DOCUMENT RESUME

ED 069 825 24 UD 013 036

AUTHOR	Tyrrell, Donald J.; And Others
TITLE	Cognitive Abilities in Culturally Disadvantaged Preschool Children. Final Report.
INSTITUTION	Franklin and Marshall Coll., Lancaster, Pa.
SPONS AGENCY	National Center for Educational Research and Development (DHEW/OE), Washington, D.C.
BUREAU NO	BR-0-B-044
PUB DATE	Sep 72
GRANT	OEG-2-700029(509)
NOTE	120p.
EDRS PRICE	MF-$0.65 HC-$6.58
DESCRIPTORS	*Cognitive Ability; *Cognitive Tests; Culturally Disadvantaged; *Disadvantaged Youth; Discrimination Learning; Lower Class; Memorizing; Pattern Recognition; Perception; *Preschool Children; Shift Studies; *Socioeconomic Status; Verbal Learning
IDENTIFIERS	Pennsylvania

ABSTRACT

Three experiments compared cognitive processes in Caucasian nursery school children from different socioeconomic backgrounds who were equated for performance on the Peabody Picture Vocabulary Test. The first demonstrated that children from lower socioeconomic backgrounds experience more difficulty in solving a series of extradimensional and intradimensional shifts than children from high socioeconomic backgrounds. This deficiency was postulated to result from a decreased rate of attentional response acquisition which is offset with experience. The second experiment demonstrated that low socioeconomic status children organize their recall of categorized and non-categorized lists to the same extent as high socioeconomic status children but are less effective in filtering inappropriate items from recall output. The final experiment demonstrated that children could identify stimuli presented via the visual and tactual modes and could make cross modal judgments regarding the equivalence of these stimuli. There were no effects of socioeconomic status on this ability, and extramodal are easier than intramodal judgments. Overall, the results do not support a simple notion of a deficiency of complex cognitive processes in lower class children. (Author/JM)

O-B-044
PA 24

Final Report

Project No. O-B-044
Grant No. OEG-2-700029(509)

Donald J. Tyrrell
Kenneth H. Brookshire
C. Richard Puff
Department of Psychology
Franklin and Marshall College
Lancaster, Pennsylvania 17604

COGNITIVE ABILITIES IN CULTURALLY DISADVANTAGED PRESCHOOL CHILDREN

September 1972

U.S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE

Office of Education

National Center for Educational Research and Development
Region III

Abstract

Three experiments compared cognitive processes in Caucasian nursery school children from different socioeconomic backgrounds who were equated for performance on the Peabody Picture Vocabulary Test. The first demonstrated that children from lower socioeconomic backgrounds experience more difficulty in solving a series of extradimensional and intradimensional shifts than children from high socioeconomic backgrounds. This defieiency was postulated to result from a decreased rate of attentional response acquisition which is offset with experience. The second experiment demonstrated that low socioeconomic status children organize their recall of categorized and non-categorized lists to the same extent as high socioeconomic status children but are less effective in filtering inappropriate items from recall output. The final experiment demonstrated that children could identify stimuli presented via the visual and tactual modes and could make cross modal judgments regarding the equivalence of these stimuli. There were no effects of socioeconomic status on this ability and extramodal are easier than intramodal judgements. The results of each experiment are related to a body of relevant theoretical and empirical literature. Overall, the results do not support a simple notion of a deficiency of complex cognitive processes in lower class children.

Final Report

Project No. 0-B-044
Grant No. OEG-2-700029(509)

Cognitive Abilities in Culturally Disadvantaged Preschool Children

Donald J. Tyrrell
Kenneth H. Brookshire
C. Richard Puff

Franklin and Marshall College
Lancaster, Pennsylvania 17604

September 1972

The research reported herein was performed pursuant to a grant with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
National Center for Educational Research and Development

Preface

The authors are deeply grateful for the welcome and pleasant cooperation extended by all school personnel during our repeated interference with normal classroom activities. In particular, we would like to thank Mrs. Eleanor O'Neil and the staff of the First Presbyterian Church Nursery School, Mrs. Esther Heller and the staff of the nursery school of St. Peter's United Church of Christ, Mrs. Dorothy Kroehler and her teachers in the Federal Head Start Program of Lancaster, and Mr. Robert Lehr, Principal, and the teachers of the Head Start classes in the Reigart and Eicholtz Schools in the city of Lancaster. We are indebted to our students, Gerald Balzano, Nancy Adams and Marge Kally, for their expertise in data reduction and analysis, and to Mr. Dick Steere and Mr. Gerald Stroud for the apparatus construction. Finally, to Mrs. Durell Neely and Mrs. Phyllis Stroud, for their monumental and dedicated effort in collecting the data, our hearty congratulations and most sincere thanks for an excellent job.

A brief version of Experiment One was presented at the 1972 meetings of the Eastern Psychological Association in Boston, Massachusetts.

Table of Contents

List of Tables

Introduction

A voluminous literature has arisen in the past decade documenting the fact that culturally disadvantaged children have developed an intellectual deficit by the time they enter elementary school, and that this deficit is difficult to eradicate through remedial programs within the public school system. It has been assumed that preventive or "interventive" educational programs, in the preschool years, have a better prognosis than later remedial programs. In general, of course, this point of view is consistent both with the concept of developmental stages of learning (e.g., Inhelder & Piaget, 1958) and with what we know about transfer of training (Shepp & Turrisi, 1966).

Recently, a number of special educational curricula or programs for preschool-aged disadvantaged children have been developed and employed. It is possible to divide these into four categories of approach.

Environmental enrichment through sensory and manipulative experience.--This approach, advocated by Hunt (1964), assumes that the disadvantaged child has experienced a deprivation of concrete, nonverbal experience which is a necessary prerequisite for the development of language and thought. The program is designed merely to remove this deficit by giving the child the opportunity to encounter a wide variety of new stimuli.

Verbal bombardment.--This approach, a commonly-used one (Gordon & Wilkerson, 1966, p. 50) has been endorsed by Jensen (in Hartrup & Smothergill, 1967, p. 134) and Weikart (1964). The assumption is that language is the basis for most learning and thinking, and since mastery of the spoken language normally is acquired between the ages of 1 and 5 (Leeper et al., 1968), it follows that the disadvantaged child should be placed in a verbally stimulating environment during the preschool years if the deficit is to be prevented or ameliorated.

Development of grammatico-logical language skills.--This approach is best exemplified by the preschool program developed by Bereiter and Engelmann (1966). The program assumes that the intellectual deficit found in disadvantaged children is not merely the result of a vocabulary deficit but involves the syntactic and structural aspects of the language. The Bereiter-Engelmann program calls for the statement of clear educational objectives in language usage which are then sought by a formal teaching program involving specific lesson units, practice, and feedback.

Stepwise perceptual discrimination training.--Deutsch (in Passow, 1963) has argued that while the disadvantaged child displays a nonfunctional language system (and therefore deficiencies in thinking and concept formation), the growth of language skill depends upon adequate perceptual development. He therefore advocates that preschool education concentrate heavily on perceptual discrimination training. This approach is consistent with the methods espoused in 1912 by Montessori (translated 1964) and, more recently, Fowler (1962) where children are exposed to stimulus discrimination problems of gradually increasing complexity on the hypothesis that acquisition of more abstract cognitive skills can develop readily out of this simpler learning. Disadvantaged preschool children do show deficits in both visual and auditory discrimination learning ability (Klaus & Gray, 1968), and programs which have employed special perceptual discrimination training have been able to report some success in outcome.

While it is too early to assess the relative strengths and weaknesses of each of these approaches to the problem of the amelioration of intellectual deficit in disadvantaged children, it is evident that the assumptions which have been made regarding the nature of the deficit are limited in scope and have not included other processes involved in learning and thinking which are well known and which have been intensively studied in other subject populations. Moreover, Jensen (in Deutsch, Katz & Jensen, 1968) has argued that the orientation of existing programs may not be the most efficaceous way of making educational progress. He states:

> "The specific basic questions that need to be answered at this stage call for carefully designed experiments rather than massive testing programs, the gathering of enormous descriptive statistics, or large scale field demonstrations ... A much more fine-grained experimental analysis of social-class differences in abilities is needed if we are to discover the specific environmental variables controlling the development of learning abilities. Then we can institute those procedures that will most effectively raise the educational potential of lower-class children, without the wasteful inefficiency of the well-meaning, but haphazard, shotgun approach that has characterized so much of the educational effort in the field so far."

The purpose of the present research program was to initiate such an experimental analysis in three areas, in order to determine whether intellectual deficits stemming from other processes of learning and thinking, processes little explored in the preschool aged disadvantaged child, can be identified and their causes traced. The "probe" areas are: (1) attentional mechanisms, (2) organizational mechanisms in memory, and (3) cross-modal transfer. None of these processes requires a language base, yet all involve processes of a higher order than perceptual discrimination learning. Together, they represent a set of mediational processes which have been virtually ignored by those working with preschool disadvantaged children, which have a firm theoretical base in the psychology of learning, and which are not unreasonable sources for the intellectual deficit found in disadvantaged children given the "content-centered", "physical", and "intellectually passive" world in which they are reared. Moreover, corrective programs for these process deficiencies would be both simple and obvious.

Since it has been demonstrated that intellectual ability is correlated with both social class (e.g., Whitman & Deutsch, 1968) and performance on tasks similar to those in the present experiments (e.g., Zeaman and House, 1963), it was deemed necessary to prevent the possible compounding of effects of social class and intellectual ability. Consequently, the general plan was to form experimental groups which differed in social class yet simultaneously control for intelligence by equating all groups for performance on a standardized test of intelligence. The Peabody Picture Vocabulary Test (PPVT) was selected as an appropriate test for the target age group and was administered to all children enrolled in one private church-related nursery school and in the two public school operated Head Start programs for two consecutive years and in selected classrooms of OEO funded Head Start programs and another private church-related nursery school for one year. Every child in each classroom was given the PPVT by one of two female experimenters each of whom was also involved in some part of the data collection for the experiments. In addition to the PPVT scores, the parental occupation and educational level of each child was also obtained. These data provided the basis for determining the social class level of the family of each child by the Hollingshead Two Factor Index of Social Position (Myers & Bean, 1968). A total of 351 children were tested and provided the pool from which the groups for each experiment were formed. Each child was assigned to one of the five social class levels suggested by Myers and Bean on the basis of the Index of Social Position score. PPVT test results are presented in Table 1-1 for each social class level separately for caucasian and Negro children.

The effects of social class on the test performance was analyzed by a series of one-way analyses of variance. These indicated a significant effect in the caucasian children on IQ (F=20; df=4,253; p<.001), MA (F=21.88; df=4,253, p<.001) and PPVT raw score (F=21.41; df=4,253; p<.001) but no differences due to social class in the Negro children. The extent of relationships between social class and these three test scores in caucasian children was measured by Pearson product moment correlations and yielded significant relationships between SES and IQ (r=-.53; df=256; p<.001), SES and MA (r=-.52; df=256; p<.001), and SES and raw score (r=-.52; df=256; p<.001). These correlations indicate that as SES decrease from high to low performance on the PPVT also decreases.

Table 1-1
Means and Standard Deviations of Subject Characteristics
for the Total Population Tested

SES Level	Race	N	SES Score	CA	PPVT Raw Score	MA	IQ
I	Cauc.	27	12.04 (1.75)	55.83 (7.36)	52.55 (7.56)	6.76 (14.99)	109.59 (13.60)
	Negro	0	--	--	--	--	--
II	Cauc.	45	21.04 (2.07)	53.95 (4.90)	52.24 (8.58)	67.55 (14.73)	110.66 (13.12)
	Negro	0	--	--	--	--	--
III	Cauc.	43	32.67 (6.24)	56.51 (4.59)	54.09 (7.51)	70.51 (14.46)	110.67 (12.83)
	Negro	4	36.25 (4.87)	58.50 (3.94)	29.50 (5.02)	36.50 (9.03)	38.00 (38.10)
IV	Cauc.	58	51.89 (5.80)	55.78 (5.48)	45.28 (14.85)	58.07 (13.50)	97.81 (19.82)
	Negro	19	53.[illegible] (5.68)	52.79 (5.13)	35.79 (10.60)	43.95 (11.2[illegible])	82.05 (23.89)
V	Cauc.	85	70.11 (4.29)	55.09 (6.33)	39.07 (9.91)	47.49 (11.57)	85.87 (19.07)
	Negro	50	69.82 (4.03)	58.08 (7.44)	37.88 (9.62)	45.88 (11.00)	77.82 (25.87)

The differences in IQ between Negro and caucasian children in SES Level IV and V were tested by *t* tests and indicated a significant difference in both SES Level IV (t=2.82; df=75; $p<.05$) and SES Level V (t=2.05; df=133; $p<.05$). These results indicate that the tested IQ of Negro children is below that of causasian children of the same social class.

The subjects for each experiment were selected from this population with the restrictions that:

1) only caucasian, English speaking children were selected,
2) HI-SES subjects were from SES Levels I or II,
3) LO-SES subjects were from SES Levels IV or V,
4) all experimental groups were equated for PPVT test performance.

Experiment 1: Discrimination Transfer in Nursery School Children from High and Low Socioeconomic Backgrounds

Purpose

In 1969, Jensen postulated the existence of two genetically distinct but functionally related abilities which underlie performance on intellectual or learning tasks. Level I abilities are described as being "associative" in nature and are operationally anchored to tasks such as digit span and serial learning tasks. Level II abilities are mediational in nature, are related to the subjects' ability to engage in self initiated activity, and are operationally anchored to cross-modal transfer tasks or clustering in free recall.

The noteworthy characteristics of Jensen's notions were the postulated differences in the distribution of these abilities as a function of social class variables. Thus, Jensen postulated that while Level I abilities were equally distributed in both upper and lower social classes, the mean of the distribution of Level II abilities in children from lower class families is shifted down relative to upper class children.

Mediation models of discrimination learning similarly postulate two distinct yet functionally related response processes which must occur during solution of a discrimination problem. Zeaman and House (1963) postulate a chain of responses consisting of an "observational" or "attentional" response which is made to a stimulus dimension and an instrumental response which is directed at a specific stimulus cue. According to their model, during the solution of a discrimination problem the subject must first learn to observe or pay attention to the relevant stimulus dimension (e.g., form, color) and then instrumentally respond by selecting the correct cue (e.g., square or red) of that relevant dimension. In a similar fashion, the Kendlers (1962) have postulated the acquisition of a covert verbal response which mediates stimulus reception and response occurrence.

In the present paper it is suggested that there exists a correspondence between the two different abilities discussed by Jensen and the two response processes postulated by the various mediational models of discrimination learning. Specifically, Level I ability is assumed to correspond to the rate of instrumental or choice response learning while Level II ability would be the analogue of the attentional or verbal mediational response acquisition rate.

This hypothesized correspondence holds an advantage experimentally, since there are several operations which are known to alter the rate of discrimination learning which are attributable to the effect of the mediational (attentional) mechanism. Perhaps the best known and replicable of these are the discrimination transfer operations of intradimensional (ID) versus extradimensional (ED) shifts. The ease of learning the second of two discrimination problems depends to a large extent upon the consistency of the relevant dimension over the two problems. Thus, if the subject experiences two form-relevant problems with different forms in each problem (an ID shift), he will acquire the second problem much more rapidly than if the relevant dimension is switched from form in Problem 1 to color in Problem 2 (an ED shift).

If there is a correspondence between the attentional (mediational) response of Zeaman and House and the Level II (mediational) ability of Jensen, and if Jensen's assumption that the mean of the distribution of Level II ability in lower class children is below that of higher class children, then a comparison of intradimensional and extradimensional shift effects should produce differential results as a function of the social class of the children.

The design appropriate to test these notions is a 2 (social class) x 2 (type of shift) analysis of variance. Thus, in the present experiment two groups of nursery school children differentiated by the measurable social class of their families were given discrimination training. One half of the children in each social class group then experienced an ID shift with the remainder experiencing an ED shift.

Finally, to assess the extent to which any differences may persist over a series of such shifts--or alternatively, to assess the rate at which children learn-to-learn successive ID or ED shifts--the experiment was expanded to a 2 (social class) x 2 (type of shift) x 11 (shift problems), with each child experiencing a series of 11 transfer problems following the original learning.

Method

Approximately 200 children attending either a private church-related nursery school or a federally-funded Head Statt nursery school in Lancaster, Pennsylvania, were administered the Peabody Picture Vocabulary Test (PPVT) and the Hollingshead Two-factor Index of Social Position (Myers & Sean, 1968). From this total pool, 32 caucasian, English speaking children were selected and assigned to one of 4 groups which were equated on MA, IQ, CA, and PPVT raw score. Two of these groups were high on the Index of Social Class (HI SES), while the other two were low (LO SES). Table 2-1 presents the relevant statistics for all groups. An analysis of variance demonstrated significant social class differences (see Appendices A; B; C; D; E) but no differences between the groups on any other measure approached significance.

A portable version of a WGTA modified for children was employed. The essential features of this apparatus consist of a one-way screen interposed between S and E, and a stimulus tray with two 3 inch

Table 2-1
Means and Standard Deviations of Subject Characteristics in Each Experimental Condition

SES Level	Type of Shift	N	SES Index	CA	MA	IQ	PPVT Raw Score
High	ID	8	17.9 (4.1)	54.1 (4.1)	62.7 (18.1)	104.0 (17.8)	49.5 (9.7)
	ED	8	20.6 (8.0)	53.7 (8.6)	58.5 (16.6)	101.5 (15.8)	46.0 (11.4)
Low	ID	8	62.0 (13.5)	53.7 (5.9)	56.7 (11.6)	100.0 (14.5)	46.4 (8.1)
	ED	8	64.2 (9.2)	54.3 (6.0)	52.0 (13.2)	93.9 (23.1)	42.5 (10.8)

recessed reward wells 8 inches apart which can be slid to S's side of the apparatus. Stimulus objects were placed over the reward wells when the tray was on E's side of the apparatus and not visible to S. A trial was initiated by sliding the tray to S's side of the apparatus. The base of the objects completely covered the reward wells and prevented S from seeing the reward object (M & M candy) hidden under the correct stimulus object. S responded by sliding the object of his choice back to reveal the reward well. A correct response uncovered a well with an M & M candy which S could place in a bag provided by E.

The stimuli were 2 inches high, ¼ inch thick masonite forms mounted upright on 4 inch square masonite bases. The stimulus pool consisted of six forms--square, circle, T, diamond, and triangle--in each of six colors--red, blue, green, white, yellow and black.

All Ss were tested individually during every school day. Each child was administered 25 two-choice trials every day with each new problem beginning on the day following criterion performance on the preceding problem. This process continued until all children had been administered the original problem and 11 successive transfer problems. Each problem consisted of one relevant dimension (either form or color) and one variable-within-trials irrelevant dimension (either color or form). Thus, on each trial two stimuli were presented which differed in both color and form with the color-form association changing randomly over trials (e.g., trial N may have included a red square versus a blue circle, while trial N+1 whould have been changed to a blue square and a red circle). On a form relevant problem, one of the forms (e.g., square) would be consistently associated with reward while on a color-relevant problem the reward was associated with one of the colors. The left-right position of the positive cue was determined by a portion of the Gellermann (1933) series and the color-form association of the positive cue on each trial was determined by another independent Gellermann series. The subject was instructed to push aside the object which he thought was hiding the candy prior to problem one and received 4 practice trials with objects which were not included in the stimulus pool.

Each problem was continued until the S was performing at a criterion level of 20 correct trials in a daily session of 25 trials. If criterion level was not attained within 150 trials (six days), a special training technique was administered. On the day following attainment of the failure criterion the subject was shown the two positive stimuli (e.g., blue square and red square) and was told that the candy would always be hidden under one of them and that he should pick those to find the candy. S was then shown the two negative stimuli (e.g., blue circle and red circle) and was told the candy would never appear under objects like those and he should not choose them. Following this training, the regular daily run of 25 trials was begun. Nine subjects within each social class were unable to solve the original learning problem without benefit of this training. This is an indication that this type of problem is extremely difficult for nursery school children.

One group of children of each social class level was given an original training problem and 11 transfer problems all of which had the same relevant dimension (intradimensional or ID shifts), while a second group of Ss from each social class level received an original training problem and 11 transfer problems in which the relevant dimensions alternated between color and form with every problem (extradimensional or ED shifts). One half the Ss in each of the four groups received a form relevant problem on original learning while the remainder were presented a color-relevant problem.

This procedure results in a 2 x 2 factorial design with two levels of social class (HI SES and LO SES) with two types of shifts (ID and ED) as the major independent variables with all groups matched on chronological age, sex, and test scores. Finally, the relevant dimensions over problems and the specific cues of each dimension were counterbalanced within each group so that there could be no systematic effects attributable to stimulus characteristics. In addition, the stimulus pairs for each problem for each S were selected from the pool so that no stimulus pair would appear on two consecutive problems and no color-form combination would appear in more than one problem.

Results

The major dependent variables were trials and errors to criterion within each problem. Analysis of both produced identical conclusions. Therefore, discussion will focus upon the errors to criterion measures. Unfortunately, complete data are not available for two Ss in the LO-SES ED group. These Ss had completed only six transfer problems at the termination of the school year. This fact complicated some of the analyses and will be discussed in detail in the appropriate results section. The mean errors to criterion for each group over all problems is presented in Table 2-2.

A 2 (social class) by 2 (type of shift) analysis of variance was computed on the errors to criterion on the original learning (see Appendix F). No significant effect of either social class or shift type was evident indicating that there were no differences in the rate of acquiring the original learning problem between the groups differing in social class or between those groups which were to receive ID or ED shifts.

Analysis of the rate of acquisition of the first transfer problem was provided by a 2 (social class) x 2 (type of shift) analysis of variance of the total errors to criterion (see Appendix G). These data indicate a significant effect due to type of shift (F=10.79; df=1,28, p<.005) with ID shifts being learned with fewer errors than ED shifts, and a significant social class by shift interaction (F=4.84; df=1,28; p<.05).

Table 2-2
Mean Errors to Criterion on Each Problem for all Experimental Groups

SES Level	Type of Shift	Orig. Learn.	Shift Problems 1	2	3	4	5	6	7	8	9	10	11
High	ID	51.9	7.4	6.8	2	1	1.6	2.1	1.7	3.5	2.3	1.3	2
	ED	68.1	15.5	10.9	16.1	19.8	9.8	1.3	11.8	3.8	2	2.4	2.5
Low	ID	46.8	3.2	19.4	3.9	1.3	20.2	10.1	2	1.3	2.2	1.3	1.3
	ED	58.1	44.4	32.1	25.9	27.8	21.5	9.0	15.3	12.8	5.8	1.8	3.6

Specific group comparisons were provided by Scheffé tests which indicated that the LO-SES ED shift group made significantly more errors than any of the other groups.

Another indication of the transfer from original learning is the savings, in errors to criterion, from original learning to the first transfer problem. These savings scores are presented in Table 2-3 along with the results of *t* tests for related scores for the differences in each group. Significant positive transfer occurred in all groups except the LO-SES ED shift group. This is another indication of the poor performance level of these subjects on the transfer problem.

Since complete data exists for all subjects on the first six transfer problems, these data were submitted to a 2 (social class) x 2 (type of shift) x 6 (transfer problems) with repeated measures over problems (see Appendix H). This analysis demonstrated significant main effects of social class (F=4.69; df=1,28; p<.05) and type of shift (F=7.30, df=1,28; p<.025) with ID shifts producing less errors than ED shifts and HI-SES subjects performing with less errors than LO-SES children. There were no problem effects but a significant shift by problems interaction appeared (F=2.46; df=5,140; p<.05). This interaction reflects the convergence of the ED and ID shift errors over successive problems. Since an F_{max} test indicated heterogeneity of variance, the analysis of variance was recomputed with a log(X+1) transformation (see Appendix I) and indicated significant main effects of social class (F=7.05; df=1,28; p<.025), type of shift (F=8.57, df=1,28; p<.01) and problems (F=3.28; df=5,140; p<.01), as well as a significant interaction of type of shift and problems (F=3.04; df=5,140; p<.01). Thus, the analysis of the transformed scores indicated a significant problems effect in addition to the significant effects uncovered in the raw data analysis.

To complete the analyses of the entire sequence of 12 transfer problems the meau errors of the six Ss who completed all problems in the LO-SES ED condition was calculated for problems 8-12 (transfer problems 7-11). The mean for each problem was then assigned to the missing cells for the two subjects who did not complete the entire sequence. A 2 (social class) by 2 (type of shift) by 11 (problems) analysis of variance was computed on these data (see Appendix J). This analysis indicated significant main effects due to social class (F=4.72; df=1,28; p<.05), type of shift (F=9.67, df=1,28; p<.005) and problems (F=4.40; df=10, 280; p<.005) as well as a significant interactions of type of shift with problems (F=2.59; df=10,280; p<.005). Thus, HI-SES children made less errors overall than did LO-SES subjects, the ID shifts were easier than the ED shifts, total errors decreased over problems and the dfference in error rate between the two shift types decreased over problems. An F_{max} test indicated heterogeneity of variance existed and a second analysis was calculated on a log(X+1) transform of the raw data (see Appendix K. The analysis produced the same pattern with effects due to

Table 2-3
Mean Savings Score and *t* for Errors on Original Learning
Minus Errors on the First Transfer Problem

SES Level	Shift Type	Savings	*t* for related distribution	df	p value
High	ID	+44.50	3.87	7	<.01
	ED	+52.60	3.47	7	<.02
Low	ID	+43.50	3.61	7	<.01
	ED	+13.75	.80	7	>.10

to social class (F=5.09; df=1,28; p<.05) type of shift (F=12.16; df=1,28; p<.001) and problems (F=5.38; df=10,280; p<.001) and the significant interaction of shift type and problems (F=2.31; df=10,280; p<.025).

One final analysis on transfer error was calculated on the error data from the last transfer problem. This analysis, a 2 (social class) x 2 (type of shift) analysis of variance (see Appendix L) indicated a significant main effect due to type of shift (F=3.07; df=1,28; p<.01) but no effect of social class and no interaction. This indicates that on this final transfer problem ID shifts were easier than ED shifts but the distinction between the children from different social classes was no longer signfficant.

Discussion

The results indicate that although there are no differences in the rate of original learning as a function of social class, performance on the transfer problems is better for HI-SES children relative to LO-SES children. In addition, ID shifts are easier than ED shifts. This differential difficulty of ID relative to ED shifts continues to exist through problem 12, whereas by problem 12 there were no effects due to the social class of the children. Finally, there was a significant decrement in overall error rate as Ss experienced the sequence of problems.

Initially, a correspondence was postulated between the two levels of ability discussed by Jensen and the two links in the response chain hypothesized by Zeaman and House. Specifically, Level I abilities were assumed to be reflected by the instrumental response acquisition rate and Level II abilities were related to the rate of acquisition of the attentional (or mediational) response.

Given those correspondences and the hypothesized differential distribution of Level I and II abilities in HI-SES and LO-SES populations, it is clear that the rate of instrumental response acquisition should not differ between populations differing in social class. However, the rate of acquisition of the appropriate attentional response should be higher in children from HI-SES relative to LO-SES backgrounds.

Since the rate of acquisition of the appropriate attentional response has been shown to affect the rate of discrimination learning and transfer, it was expected that children from differing social class backgrounds would differ in the rate at which they acquire and transfer discriminative responses. This difference in rate should appear on original learning as a main effect of social class. The analyses did not indicate such a main effect. However, it is possible that any differences in learning rate were masked by the difficulty level of the problem. Recall that approximately 60% of the subjects in both social class groups failed to learn the original problem without the special training technique. It is conceivable that social class differences in acquisition were indeed eliminated by that special technique. However, the fact that there were equal numbers of children in each social class condition who received that special training argues strongly against social class differences in initial learning rate.

The logic is not as straightforward for predicting an outcome on the first transfer problem. At the conclusion of Problem 1 all children will have acquired the appropriate attentional (mediational) response as well as the correct instrumental response. All of the children then experienced a second problem in which there were totally new cues of both the form and color dimension. This means that all Ss had to learn a new instrumeatal response and this process should have no differential between-group effects. One half of the children in each social class group experienced an ID shift with the remainder experiencing an ED shift. Since the ID shift groups have already acquired the attentional response appropriate for the shift problem (and if they transfer that response across problems), they should learn the second problem rather rapidly. However, those subjects experiencing the ED shift will have to extinguish this original attentional response and acquire a novel one (indeed, one which has undergone extinction on Problem 1). To the extent that lower class children have a decreased rate of acquiring this new attentional response, they will experience more difficulty than the upper class children. This difference is reflected in the social class by shift interaction with the lower class ED group performing below the level of other groups. Not only was that interaction significant but the lower class ED shift children experienced no significant positive transfer from Problem 1 to Problem 2 whereas children in all other groups did experience such positive transfer.

If the lower mean of Level II abilities in lower class children is relatively permanent due to its genetic or physiological basis (as postulated by Jensen), then the poorer performance of these children witnessed in the first transfer problem should continue throughout the series of transfer problems. The analysis of the first six transfer problems as well as the entire sequence of 11 transfer problems produced a significant main effect of social class with the lower class children performing at a poorer level than the high class children. These data are consistent with expectations based upon Jensen's reasoning. However, the picture is muddied by the fact that the social class difference occurs as a main effect and not as an interaction with problem type and by the absence of a significant difference due to SES on the final problem. The former indicates that the lower class children were inferior on both ID and ED shifts. This would implicate a deficiency in some process other than the rate of attentional response acquisition. Alternative processes in which the deficiency may exist are the instrumental response acquisition rate or some parameter controlling transfer of the attentional response between problems. Although experimental operations do exist to assess the effect of both of these processes, the present design did not provide the conditions necessary for such analyses.

The final analysis, that of errors on the last transfer problem, indicates that whatever deficiencies are coincident with lower socioeconomic status are overcome by continued training and are not permanent as suggested by Jensen. Performance on the final problem was not affected by the social class factor although the type of shift was still important.

In summary, children from families of lower socio-economic status do suffer from a deficiency in the acquisition and transfer of discriminative responding and this deficiency appears to be associated with an inferred cognitive (mediational) process such as selective attention. Contrary to the expectancies of Jensen, the deficiency does not seem to be permanent and, therefore, neither constitutional nor genetic. In fact, with relatively short term exposure to several similar problems the mediational deficiency disappears and the lower class children begin responding at a level equivalent to the higher class children.

It is of interest that this deficiency disappears without specific training in mediational responding. At no time during the experiment were the children given any instructions in how to solve the individual problem or in transferring the principle of solution across problems. Thus, to the extent that the mediational deficiency was overcome, it occurred spontaneously and without formal instruction.

Experiment 2: Free Recall of Categorized and Non-Categorized Material in Nursery School Children as a Function of Social Class

Purpose

Jensen (1969, 1971, etc.) has offered a major attempt to specify the nature of the cognitive deficits of culturally deprived children. Briefly, he has drawn a distinction between what he has called Level I and Level II abilities. Level I ability is conceived of as involving a basic capability for associating things, for example, in a simple S-R type of paradigm, or task. This kind of ability is further thought to be equally distributed among differing SES groups, and perhaps to develop prior to Level II abilities. Level II abilities are thought to comprise more complex kinds of cognitive capabilities such as those presumed to be required in conceptual learning, problem solving, and other kinds of tasks which are heavily dependent upon verbal mediation. It is further proposed that Level II abilities are not equally distributed within different SES groups, but are distributed about a higher mean among higher SES children. Level II abilities are either critically dependent upon the prior existence of sound Level I abilities, or simply appear later in the developmental sequence.

In short, Jensen has proposed that culturally deprived, or low SES, children suffer a cognitive deficit in the capacity for complex conceptual activities, but have no impairment of the more basic associative ability.

In his efforts to support this view, Jensen has emphasized the results of several studies which have used the free-recall learning and memory paradigm. This emphasis is justified on the basis of his assumption that this task can be an instance of either a Level I or a Level II type of task depending upon the precise nature of the stimulus materials which are presented, thus affording an otherwise high degree of commonality between the tasks at the two levels. More specifically, Jensen asserts that when the stimulus list comprises items which are "unrelated", in the sense that they cannot be readily classed into a few common conceptual categories, then only Level I abilities are required for its recall or mastery, and differences are thus not to be expected between children from different SES backgrounds at any stage of development. On the other hand, if such categories are built into the list the task becomes, according to Jensen, one which requires Level II abilities. Consequently, superior performance is expected in recall of a categorized list by high SES children (if they are far enough along in the developmental sequence).

Jensen (1969, 1971) has concluded from his work that there are, in fact, no differences in the amount recalled from non-categorized stimulus lists by HI-SES and LO-SES groups of children.

The research which leads to this conclusion has not been terribly explicitly cited by Jensen, but presumably must come from any or all of three specific studies. One of these is a study by Jensen and Frederiksen (see Jensen, 1971) which is as yet unpublished and has not been sufficiently described to permit any adequate evaluation. However, it has been noted that it involved confounding of SES and race. A second somewhat relevant study (Jensen, 1968) investigated digit span performance (rather than recall of semantic material) in children with CAs between three and five years. No significant differences were found in performance by children from families on public assistance and those attending private nursery schools. Once again there was substantial confounding of SES and race. A third study can perhaps be drawn upon if the specification of social class is relaxed somewhat. Jensen (1961) describes comparisons of recall of a sequence of familiar, but unrelated, objects by Mexican-American and Anglo-American children. Jensen noted that the two groups did not differ with respect to standard SES indicators such as father's occupation, but there were differences between the groups in IQ and scholastic achievement which would normally be associated with differences in SES. The results in a first experiment with fourth grade children showed that when the two groups were matches at an IQ of about 80, the Mexican-American children actually recalled more items, but that when the groups were matched at an IQ of about 115 there were no significant differences. A second experiment with sixth grade children also revealed no significant differences between these two subject groups.

Two studies have concerned the influence of SES level upon free recall performance with categorized lists. One of these is an unpublished dissertation by Glasman (1968) and the other is the study by Jensen and Frederiksen. Jensen (1971) reports that both were similar methodologically, i.e., they both involved confounding of SES and race, stimulus lists comprising four categories of five objects, and a sequence of five trials. Close agreement was also observed with regard to the important features of the results. Comparisons of the amount recalled and the degree of categorical organization (clustering) revealed no differences between HI-SES and LO-SES groups at the kindergarten level. However, with children in the fifth grade, the HI-SES Ss were substantially superior in both recall and clustering. Jensen also offered the additional important observation that while the LO-SES children did not demonstrate much clustering on the basis of the common taxonomic categories, they did produce many pairs of items that were related on the basis of some functional association and which were idiosyncratic. If there is formal evidence in support of this it must be the Jensen and Frederiksen paper.

Thus, the available evidence is at least consistent with the conclusion which Jensen has drawn. However, a number of important problems still need to be resolved.

The most general problem is the fact that Jensen has derived his theoretical notions, i.e., about the kinds of abilities which are required by different types of lists, from the results of the studies cited here. To use these notions to <u>explain</u> the same data represents, of course, a very obvious instance of circularity. Equal success with new data is required to demonstrate that these theoretical notions have any general explanatory value.

Secondly, it seems almost inconceivable, given his position on the genetic issue, that Jensen continues to base his conclusions about the nature of the effects of cultural deprivation upon evidence from studies where SES and race are totally or substantially confounded. Clearly, these variables have to be separated before any conclusions about SES effects can be made with confidence.

Finally, Jensen's conclusions rest heavily upon his assumptions and his evidence concerning the nature of organizational processes in learning and memory. For example, he asserts that categorized lists require Level II abilities because they involve substantial transformation by <u>S</u>s of the randomized input order, and that this is not the case with non-categorized lists. He further suggests that <u>S</u>s at different SES levels used different modes of organization, and that not much organization by any <u>S</u>s with any type of list takes place at the kindergarten level. The proper evaluation of these kinds of conclusions would appear to warrant a much more detailed examination of organization in recall than has yet been provided.

Thus, the purposes of the present research were to correct some of the weaknesses of previous research in this area, to further assess the utility of Jensen's analysis, and to extend our understanding of the cognitive and conceptual consequences of cultural deprivation.

<u>Method</u>

The stimulus items were colored picture vocabulary cards from Level #1 of the Peabody Language Development Kits. There were two types of lists, categorized (C) and non-categorized (NC), and there were also two separate lists of each type for purposes of internal replication and control.

As shown in Appendix M, the C lists comprised three items from each of four taxonomic categories for a total list length of 12 items. More specifically, C list No. 1 included categories of things to eat, vehicles, animals, and articles of clothing, while the categories of C list No. 2 were fruits, articles of furniture, tools, and people representing familiar occupations.

The NC lists comprised 12 items chosen so as to avoid more than single instances of the kinds of broad taxonomic categories built into the C lists.

Duplication of initial phonemes was avoided within all lists. Additionally, C list No. 1 and NC list No. 1 were matched in the sense that one item from each of the categories in C list No. 1 also appeared in NC list No. 1. Similarly, one item from each of the categories in C list No. 2 also appeared in NC list No. 2. Subsequent reference will be made to list pairs 1 and 2.

The apparatus included the actual picture cards from the kit which were manually presented. A small flashing light, seen only by E, was used to pace the presentation of the cards, and an Ampex Micro 24 tape recorder was used to record S's responses.

The Ss in this study were 72 Caucasian preschool children. As shown in Table 3-1, there were 18 Ss in each of the four major SES X type of list groups. Furthermore, the 18 Ss in each of these groups were divided into nine who received list No. 1 of that particular type and nine who were given list No. 2.

The two SES levels were established as follows: the high SES Ss all attended either of two private, church-related nursery schools and had Hollingshead Two-Factor Index of Social Position (Myers & Bean, 1968) values falling at Levels I and II. The low SES Ss attended either of two Head Statt programs. These Ss were all from Levels IV and V according to the Hollingshead Index. The mean SES index scores for all groups are shown in Table 3-1. The mean scores for high and low SES Ss were compared by t-tests within each of the four type of list X list No. conditions with the result that every comparison was significant at the .001 level or beyond.

The two SES levels were matched as closely as possible, within the limitations of the available S population, for MA, IQ, CA, and sex. The MA and IQ scores were determined with Form A of the Peabody Picture Vocabulary Test (PPVT). The Ss for this experiment were selected from a larger number of children who were screened with the PPVT. Testing with the PPVT took place a minimum of four weeks before the conduct of the experiment.

The means and standard deviations of the matching variables, CA, MA, IQ, and PPVT raw scores, are shown for all conditions in Table 3-1. Comparisons between HI-SES and LO-SES groups by t tests revealed that none of the differences achieved the .05 level of confidence. Thus, for example, the HI-SES Ss given C List No. 1 were equivalent on these measures to the LO-SES Ss given the same list. The same was true for C List No. 2, and for C lists 1 and 2 combined, etc.

Procedure

The data were collected by an adult female E. This same E had previously tested about half of the Ss with the PPVT and always spent about a week in a cladsroom helping the teacher and getting acquainted before any of the children in that class were used in the experiment.

Table 3-1

Means and Standard Deviations of the Important Characteristics of the Subjects in Each Condition

SES Level	Type of List	List No.	N	M/F	SES Index	CA	MA	IQ	PPVT Raw Score
High	Cate-gorized	1	9	5/4	21.55 (8.63)	58.55 (6.27)	69.55 (9.82)	109.11 (5.64)	54.00 (4.66)
		2	9	4/5	20.33 (6.10)	55.44 (6.42)	67.66 (10.94)	109.33 (8.31)	53.00 (5.47)
		1&2	18	9/9	20.94 (7.27)	57.00 (6.36)	68.61 (10.13)	109.22 (6.89)	53.50 (4.96)
	Non-Cate-gorized	1	9	4/5	19.44 (7.46)	53.11 (5.30)	64.88 (9.07)	109.33 (6.48)	51.66 (4.82)
		2	9	5/4	21.22 (7.27)	60.22 (11.68)	65.88 (6.93)	109.77 (4.14)	46.00 (17.63)
		1&2	18	9/9	20.33 (7.21)	56.66 (9.53)	65.38 (7.85)	109.55 (5.28)	48.83 (12.87)
Low	Cate-gorized	1	9	6/3	56.00 (10.45)	54.66 (5.52)	63.55 (10.81)	107.00 (8.07)	50.88 (5.62)
		2	9	6/3	55.11 (8.62)	59.33 (4.18)	68.33 (8.50)	108.22 (7.83)	53.33 (3.90)
		1&2	18	12/6	55.55 (9.31)	57.00 (5.32)	65.94 (9.72)	107.61 (7.74)	52.11 (4.86)
	Non-Cate-gorized	1	9	7/2	56.33 (10.81)	55.77 (5.84)	64.11 (10.63)	107.11 (8.46)	51.66 (5.35)
		2	9	5/4	53.66 (9.64)	55.44 (5.68)	63.88 (9.68)	107.44 (7.45)	50.88 (4.56)
		1&2	18	12/6	55.00 (10.03)	55.61 (5.59)	64.00 (9.86)	107.28 (7.74)	51.05 (4.83)

The experimental task was explained to the Ss as follows:

> "We are going to play a game to see how well you remember pictures. I'll show you some pictures and tell you their names. When we're finished, I'll ask you what the pictures were, and you'll tell me the names of all of the ones that you can remember. There may be too many to remember all of them, so just try to remember as many as you can. Do you think that you understand the game? Do you want to ask any quwations?"

The stimulus cards were then shown and labelled by E at the rate of one every three seconds. After the final stimulus card, E turned on the concealed tape recorder with a foot-pedal and said, "Now, what pictures do you remember?" The recall period lasted until S had given no more items for a period of about 15 seconds and had been prompted for any additional items. Then S was given some mild praise and encouragement.

The procedure was repeated for a total of six trials. The order of presentation of the items on every trial was haphazard, and was determined by a shuffling of the cards before each trial. After the completion of all trials each S was given a choice of 10¢ prizes.

Results

The initial set of analyses concerned several indices of the amount of recall and its accuracy.

The first data to be examined were the mean number of items recalled correctly from the stimulus list. For this analysis extra-list intrusions were never counted as correct and responses which S repeated within a single trial were tabulated as only a single correct response. Means for all conditions are shown in Table 3-2.

The application of the F_{max} test revealed no significant heterogeneity of variance and inspection revealed no linear relationship between means and variances. The data were then analyzed according to a design comprising three between-Ss variables: high and low SES, C and NC types of lists, and lists 1 and 2 of each type. Trials represented the single within-Ss variable.

The results of the analysis are shown in detail in Appendix N. Greater correct recall was observed with the C list than with the NC list, $F\ (1,64) = 17.25$, $p<.001$. The second-order interaction involving C vs. NC lists x SES level x Lists 1 and 2 was also significant, $F\ (1,64) = 9.66$, $p<.005$. The first-order interaction of C vs. NC lists x level of SES fell just short of significance, $F\ (1,64) = 2.80$, $.10>p>.05$. Additionally, the overall effect of trials was significant, $F\ (5,320) = 50.43$, $p<.001$, as was the interaction of trials x C vs. NC lists, $F\ (5,32) = 2.48$, $p<.05$.

Table 3-2

Mean Number of Items Recalled Correctly

SES Level	Type of List	List No.	Trials 1	Trials 2	Trials 3	Trials 4	Trials 5	Trials 6	Means for: List No.	Means for: Type of List	Means for: SES Level
High	Categorized	1	5.56	6.56	7.22	7.78	7.89	9.33	7.39	6.97	6.59
		2	5.00	5.89	6.11	6.11	5.89	7.56	6.56		
		1 & 2	5.33	5.94	6.94	7.44	7.44	8.72	6.97		
	Non-Categorized	1	5.00	5.89	6.11	6.11	5.89	7.56	6.09	6.20	
		2	5.00	6.11	6.67	6.22	6.67	7.22	6.32		
		1 & 2	5.00	6.00	6.39	6.17	6.28	7.39	6.20		
Low	Categorized	1	4.44	5.89	6.78	7.33	7.89	8.67	6.83	7.47	6.57
		2	5.89	7.67	8.33	8.78	8.67	9.33	8.11		
		1 & 2	5.17	6.78	7.56	8.06	8.28	9.00	7.47		
	Non-Categorized	1	4.22	6.33	6.67	6.22	7.33	7.78	6.43	5.67	
		2	2.78	4.89	5.22	5.11	5.11	6.33	4.91		
		1 & 2	3.50	5.61	5.94	5.67	6.22	7.06	5.67		

To clarify the role of the supposedly equivalent lists, separate analyses were performed for each of the two pairs of lists. Each of these analyses involved a C and NC list, levels of SES, and trials. The results for list pair 1 indicated only a marginal effect of C vs. NC HI-SES, F (1,32) = 3.51, .10>p>.05, and a highly significant trials effect, F (5,160) = 28.12, p<.005. The results for list pair 2, on the other hand, revealed highly significant effects for C vs. NC lists, F (1,32) = 16.72, p<.001, C vs. NC x level of SES, F (1,32) = 12.37, p<.001, and trials, F (5,160) = 22.76, p<.001. Post hoc comparisons by the Scheffé method were performed at the .05 level upon the results with list pair 2. These comparisons revealed that correct recall from the C list was significantly greater than from the NC list for the LO-SES Ss, but not for the HI-SES Ss. Furthermore, HI-SES Ss recalled significantly more than LO-SES Ss in the case of the NC list, but with the C list the LO-SES Ss actually recalled significantly more items!

Thus, the results of the analysis of the number of items recalled correctly can be summarized as follows. There was no overall effect of level of SES. However, significant SES effects were observed in the case of one of the two pairs of lists where it was found that more items were recalled from the NC list by the HI-SES Ss while more were recalled from the C list by the LO-SES Ss. A significant overall effect indicating greater correct recall from the C lists was observed. It is interesting though that the greatest part of this overall effect was most specifically attributable to the performance of the LO-SES Ss given list 2.

A second analysis concerned the total number of responses produced, or emitted, by S during the recall period. The data for this analysis included all items recalled correctly, extra-list intrusions, and repetitions, or duplications, of the same item within a single trial. These data are shown in Table 3-3.

Once again there was no indication of heterogeneity of variance nor correlation of means and variances, and the analysis of variance was performed upon the raw scores. The results of this analysis are reported in detail in Appendix O.

Overall, significantly more words were produced in recall from C lists than from NC lists, F (1,64) = 5.94, p<.025. The main effect of SES was not significant, F (1,64) = 2.36, p>.10, but the first-order SES x trials interaction was significant, F (5,320) = 2.64, p = .025. The nature of this interaction was such that, after the first trial, scores for the LO-SES Ss became consistently higher than those for the HI-SES Ss, though comparisons by the Scheffé method indicated that the difference reached significance at the .05 level only for the fifth trial.

Additionally, the first-order interaction of SES x C vs. NC lists was marginally significant, F (1,64) = 3.23, .10>p>.05, as

Table 3-3

Mean Scores for Total Number of Responses Produced

SES Level	Type of List	List No.	Trials 1	Trials 2	Trials 3	Trials 4	Trials 5	Trials 6	Means for: List No.	Means for: Type of List	Means for: SES Level
High	Categorized	1	6.00	7.00	8.22	8.33	9.22	11.56	8.40	8.02	7.84
		2	5.44	6.00	7.11	8.11	8.00	11.22	7.65		
		1 & 2	5.72	6.50	7.67	8.22	8.61	11.39	8.02		
	Non-Categorized	1	5.44	7.44	7.11	8.44	8.11	11.33	7.98	7.66	
		2	5.33	6.67	8.11	7.22	7.33	9.33	7.33		
		1 & 2	5.39	7.06	7.61	7.83	7.72	10.33	7.66		
Low	Categorized	1	6.00	7.22	9.00	10.67	11.67	12.56	9.52	9.90	8.70
		2	6.22	9.44	10.44	10.22	12.78	12.56	10.28		
		1 & 2	6.11	8.33	9.72	10.44	12.22	12.56	9.90		
	Non-Categorized	1	4.89	7.44	9.00	8.22	12.67	12.11	9.06	7.51	
		2	3.56	5.89	6.00	6.56	5.78	8.00	5.96		
		1 & 2	4.22	6.67	7.50	7.39	9.22	10.06	7.51		

was the second-order interaction of SES x C vs. NC lists x lists 1 and 2, F (1,64) = 3.06, .10>p>.05. These marginal effects reflected some tendency for LO-SES Ss to produce more words in recall from C lists than HI-SES Ss, and for C lists to be superior to NC lists primarily with these LO-SES Ss. Both of these effects were somewhat stronger with the second pair of lists.

Thus, the analysis of the total number of responses produced showed essentially the same pattern as was found for correct recall. After the initial trial, the LO-SES Ss consistently produced more responses, and by the fifth trial the difference was significant. There was also, once again, the marginal tendency for LO-SES Ss to produce more words than HI-SES Ss with the C lists and for C to be superior to NC for these Ss, especially with the second lists.

The number of incorrect responses in recall was also tabulated and examined. An error was scored each time S emitted a response which was not one of the items actually presented in the stimulus list (called an extra-list intrusion) or repeated the same response more than once on the same recall trial (called a duplication). Since E supplied the "names" for the stimulus pictures any responses other than these were scored as intrusions. However, singular/plural confusions were accepted as correct. The total number of such errors is shown for each condition in Table 3-4.

The variance for the individual conditions ranged from .45 to 9.04--variation which was significant at .05 by the F-max test--so the analysis of variance was performed upon the log (X + 1)-transformed scores. The details of this analysis are shown in Appendix P.

This analysis indicated that LO-SES Ss made significantly more erroneous responses overall than did the HI-SES Ss, F (1,64) = 7.66, p<.01. There were significant overall list effects with more errors being committed with the first pair of lists than with the second, F (1,64) = 3.98, p<.05. The trials effect was highly significant, F (5,320) = 22.77, p<.001. As can be seen in Table 3-4, the significant trials effect can be attributed to the increasing number of errors during later trials. The second order interaction of SES x Type of List x List No. was significant, F (5,320) = 2.26, p<.05, as was the third order interaction of all variables, F (5,320) = 3.12, p<.01. These interactions were not specifically tested, but appear to indicate that the largest discrepancy between LO-SES and HI-SES was with the C lists (especially the C1 list) while the scores were nearly equivalent with the NC2 list. Furthermore, these effects seemed to be strongest in the middle trials because of a big jump in errors for HI-SES Ss on the last trial.

It should be noted that the raw (untransformed) scores were also analyzed with essentially the same outcome except that the interactions were only marginally significant and the third order

Table 3-4

Mean Number of Total Errors

SES Level	Type of List	List No.	Trials 1	2	3	4	5	6	Means for: List No.	Means for: Type of List	Means for: SES Level
High	Categorized	1	.44	.44	1.00	.56	1.33	3.22	1.17	1.13	1.29
		2	.33	.67	.44	1.00	1.00	3.11	1.09		
		1 & 2	.39	.56	.72	.78	1.17	3.17	1.13		
	Non-Categorized	1	.44	1.57	1.00	2.33	2.22	3.78	1.89	1.45	
		2	.33	.56	1.44	1.00	.67	2.11	1.02		
		1 & 2	.39	1.06	1.22	1.67	1.44	2.94	1.45		
Low	Categorized	1	1.56	1.33	2.22	3.33	3.78	3.89	2.69	2.43	2.13
		2	.33	1.78	2.11	1.44	4.11	3.22	2.17		
		1 & 2	.94	1.56	2.17	2.39	3.94	3.56	2.43		
	Non-Categorized	1	.56	1.11	2.33	2.00	5.33	4.33	2.61	1.83	
		2	.78	1.00	.78	1.44	.67	1.67	1.08		
		1 & 2	.67	1.06	1.56	1.72	3.00	3.00	1.83		

interaction dropped below the .10 level of confidence. These results are shown in Appendix Q.

Tables 3-5 and 3-6 show the breakdown of total errors into the separate categories of intrusions and duplications. It can be seen that duplications represent a much larger proportion of the total than do intrusions. Since so many of the individual Ss in the HI-SES group did not make any errors, the variances between conditions differed widely, and no analyses were performed upon these data.

Thus the error data showed a general increase across successive trials, reaching magnitudes in later trials representing up to between approximately 17 to 35 percent of the total words produced. This somewhat surprising trend makes more sense when it is remembered that duplication of the same response within a trial comprised the most typical kind of error. Overall, LO-SES Ss emitted a greater number of errors than did HI-SES Ss, though this must be interpreted in light of the significant third-order interaction of all of the variables.

In an effort to corroborate the general trend emerging from the previous analysis, a score was found for each S representing the proportion of the total words produced on each trial which were correct. The mean percentages for each of the conditions are shown in Table 3-7.

Overall, a higher percentage of the words produced by the HI-SES Ss were correct than was the case with the LO-SES Ss, $F\ (1,64) = 9.65$, $p<.005$. The significant trials effect, $F\ (5,320) = 10.75$, $p<.001$, indicated that the decrease over trials in the percentage of recall which was correct was highly reliable. The percent correct recall from the second pair of lists was greater than from the first, $F\ (1,64) = 4.49$, $p<.05$, though this effect interacted with type of list and trials, $F\ (5,320) = 2.84$, $p<.025$. Finally, the third order interaction of all four variables was significant, $F\ (5,320) = 2.29$, $p<.05$. All of the findings of this analysis are reported in Appendix R.

Simply as a precaution, the data were also analysed after transforming each of the values by 2.arcsin $-\sqrt{X}$. The results were quite identical except that the main effect of lists (1 vs. 2) dropped below the .05 level of significance.

The pattern here, then, was that the HI-SES Ss showed an overall higher percentage of recalled responses which was correct, and in general the percentages decreased over trials. And, of course, the significant 4-way interaction must be kept in mind.

A second major set of analyses was designed to examine the extent to which categorical organization, or clustering, was found in recall from the categorized lists. In other words, the purpose was to specify the degree to which the Ss had recalled the items in clusters of words from the same category, thereby giving evidence of having utilized the categorical structure of the stimulus materials.

Table 3-5

Mean Number of Extra-List Intrusion Errors

SES Level	Type of List	List No.	Trials						Means for:		
			1	2	3	4	5	6	List No.	Type of List	SES Level
High	Categorized	1	.22	.11	.11	.11	.22	.11	.15	.08	.09
		2	.00	.00	.00	.00	.00	.00	.00		
		1 & 2	.11	.06	.06	.06	.11	.06	.08		
	Non-Categorized	1	.00	.11	.00	.00	.00	.00	.02	.09	
		2	.11	.00	.44	.00	.22	.22	.17		
		1 & 2	.06	.06	.22	.00	.11	.11	.09		
Low	Categorized	1	.78	.44	.33	.78	1.33	.44	.68	.44	.42
		2	.00	.33	.22	.22	.22	.22	.20		
		1 & 2	.39	.39	.28	.50	.78	.33	.44		
	Non-Categorized	1	.33	.33	.22	.33	1.00	.56	.46	.39	
		2	.44	.56	.22	.22	.22	.22	.31		
		1 & 2	.39	.44	.22	.28	.61	.39	.39		

36

Table 3-6

Mean Number of Within-Trial Duplication Errors

SES Level	Type of List	List No.	Trials 1	2	3	4	5	6	Means for: List No.	Type of List	SES Level
High	Categorized	1	.22	.33	.89	.44	1.11	2.11	.85	.97	1.17
		2	.33	.67	.44	1.00	1.00	3.11	1.09		
		1 & 2	.28	.50	.67	.72	1.06	2.61	.97		
	Non-Categorized	1	.44	1.44	1.00	2.33	2.22	3.78	1.87	1.36	
		2	.22	.56	1.00	1.00	.44	1.89	.85		
		1 & 2	.33	1.00	1.00	1.67	1.33	2.83	1.36		
Low	Categorized	1	.78	.89	1.89	2.56	2.44	3.44	2.00	1.98	1.72
		2	.33	1.44	1.89	1.22	3.89	3.00	1.96		
		1 & 2	.56	1.17	1.89	1.89	3.17	3.22	1.98		
	Non-Categorized	1	.33	.78	2.11	1.67	4.33	3.78	2.17	1.45	
		2	.33	.44	.56	1.22	.44	1.44	.74		
		1 & 2	.33	.61	1.33	1.44	2.39	2.61	1.45		

Table 3-7

Correct Recall as a Percentage of Total Words Produced

SES Level	Type of List	List No.	Trials						Means for:		
			1	2	3	4	5	6	List No.	Type of List	SES Level
High	Categorized	1	90.74	94.89	89.53	95.15	87.37	82.54	90.04	90.07	88.16
		2	95.19	91.37	96.23	89.02	88.07	80.76	90.11		
		1 & 2	92.96	93.13	92.88	92.09	87.72	81.65	90.07		
	Non-Categorized	1	95.38	83.14	88.98	79.05	82.57	73.70	83.80	85.24	
		2	94.34	93.12	86.10	90.04	90.39	78.09	88.68		
		1 & 2	94.86	88.13	87.54	84.55	86.48	75.89	86.24		
Low	Categorized	1	77.61	85.24	83.40	79.15	75.03	73.23	78.95	80.85	81.21
		2	94.80	82.38	83.32	87.11	70.84	78.05	82.75		
		1 & 2	86.21	83.81	83.36	83.13	72.94	75.64	80.85		
	Non-Categorized	1	88.98	86.35	74.79	76.63	66.37	65.68	76.47	81.57	
		2	85.74	86.75	90.76	85.54	89.24	82.01	86.67		
		1 & 2	87.36	86.55	82.77	81.08	77.81	73.84	81.57		

The raw data used for the clustering analyses were the complete recall sequences. That is, intrusions and duplications were included in the analyses rather than being disregarded as if they had not occurred. This procedure was followed on the assumption that the consideration of all of the responses actually produced by the S gives the fullest reflection of the use of conceptual relationships.

The first step in the treatment of the data was to determine the observed amount of clustering. The unit of measurement was the stimulus category repetition. An observed unit of SCR, O(SCR), was scored each time the recall of a word from any category was followed immediately by the recall of another word from the same category.

By themselves, O(SCR) scores are not readily interpretable because the number of units of O(SCR) which is possible in a given sequence, or which can be expected on the basis of chance, varies with some parameters of the sequence such as the total number of words recalled, the number of categories represented in the sequence, and the distribution of recall across the categories. Several types of derived scores have been developed, each of which expresses O(SCR) relative to chance or maximum possible values. Each of the measures is purported to have various strengths and weaknesses (i.e., reflecting its degree of independence of the recall parameters). No single measure has yet achieved compelling theoretical and empirical support, so a number of the different measures were applied here. Brief definitional formulas for all measures are shown in Appendix S.

The first frame of reference for evaluating the O(SCR) scores was the number of units of SCR to be expected on the basis of chance, E(SCR), corresponding to each of the O(SCR) values. The E(SCR) values were calculated according to the procedure described by Bousfield and Bousfield (1966). Mean observed-minus-expected clustering difference-scores are shown for each condition in Table 3-8.

The extent to which the clustering within each condition deviated from a chance amount was assessed by performing t-tests for matched observations upon the difference scores for each trial and each condition. These results are shown in Table 3-9. As can be seen from the results of these comparisons, a substantial degree of clustering was present in recall by these Ss. On the first two trials 5 of the 12 comparisons were significant at the .05 level, but by the last four trials, only 2 of the 24 comparisons fell short of the .05 level. It can also be seen that the clustering by the LO-SES Ss given list 2 and for the two lists combined, represents the only occurrence of significant deviation from chance on the first trial. Overall, the LO-SES Ss seem to have imposed at least as much clustering as did the HI-SES Ss, and perhaps even began at a higher level in early trials.

A more direct comparison between conditions was provided by entering the O-E(SCR) difference scores into an analysis of

Table 3-8

Mean Observed-Expected SCR Difference Scores for Categorized Lists

SES Level	List No.	Trials						Means for:	
		1	2	3	4	5	6	List No.	SES Level
High	1	.42	.42	1.45	1.93	1.39	2.51	1.35	1.21
	2	.28	.73	1.25	1.59	1.10	1.49	1.07	
	1 & 2	.35	.57	1.35	1.76	1.25	2.00	1.21	
Low	1	.22	.35	1.71	1.76	1.59	2.19	1.30	1.66
	2	.85	1.53	1.83	2.65	2.61	2.61	2.01	
	1 & 2	.53	.94	1.77	2.20	2.10	2.40	1.66	

Table 3-9

Results of *t*-tests Between Observed and Expected Clustering (SCR) Scores in the Categorized Lists

SES Level	List No.	Results	Trials 1	Trials 2	Trials 3	Trials 4	Trials 5	Trials 6	df
High	1	*t*	1.36	1.78	2.84	5.01	5.60	4.90	[illegible]
		p			*	**	***	**	
	2	*t*	.68	1.83	4.46	3.45	1.73	2.07	8
		p			**	**		†	
	1 & 2	*t*	1.39	2.51	4.76	5.95	3.74	4.48	17
		p		*	***	***	**	***	
Low	1	*t*	.57	.83	2.96	3.32	2.34	3.77	8
		p			*	*	*	**	
	2	*t*	3.21	3.26	3.49	5.41	5.20	7.66	8
		p	*	*	**	***	***	***	
	1 & 2	*t*	2.26	2.78	4.67	6.01	4.89	7.26	17
		p	*	*	***	***	***	***	

† $.10 > p > .05$
* $p < .05$
** $p < .01$
*** $p < .001$

variance. The design for this analysis involved SES (HI vs. LO), List No. (1 vs. 2) and Trials (1-6). As shown in detail in Appendix T, the analysis of these scores revealed no significant differences except for the trials effect.

Thus, while the LO-SES Ss may, in some instances, have shown significant deviations from chance earlier, the overall analysis failed to reveal any reliable differences between the magnitudes of the O-E(SCR) scores for the two SES levels.

A second kind of clustering score was formed by taking each S's O(SCR) value on each trial as a proportion of the maximum possible O(SCR) value. This kind of index of clustering has previously been used by Puff (1970) and others. As shown in Appendix S, the maximum possible value is given by taking the total recall minus the number of categories represented in recall. This means that if S recalls only a single word from each category the maximum possible value will be zero, and so will the observed value, giving an indeterminate score. It was reasoned that the best procedure in cases like this was to exclude such a S from the analysis. Because of the repeated measurement nature of the design, when a S had a single such score, his data for all trials had to be dropped. Accordingly, one S was dropped from each of the conditions of : HI-SES, List 1; HI-SES, List 2, LO-SES, List 1.

The resulting mean ratio scores are shown in Table 3-10. It can be seen that the average over all Ss on the first trial indicates a level of clustering which is about 53% of the maximum possible amount.

The analysis of these scores is shown in Appendix U. It can be seen that the only significant source of variability is that for trials, F (5,145) = 3.45, $p<.005$. Specific tests were not conducted, but it can be seen from the results in Table 3-9 that the trend was for scores to reach a peak in the middle trials and then to decline somewhat again.

A clustering score presented by Roenker, Thompson, and Brown (1971), called the adjusted ratio of clustering (ARC) score, was also found for the present data. As shown in Appendix S, the ARC score expresses the O-E(SCR) deviation for a S relative to the maximum possible deviation of O(SCR) from E(SCR). As was the case with the simple ratio scores, ARC scores are also indeterminate when S recalls only a single word from each category, and the same three Ss were excluded from this analysis. The mean ARC scores for each of the conditions are presented in Table 3-11 and the analysis of these scores is reported in Appendix V. It shows exactly the same pattern with only a significant trials effect, F (5,145) = 4.54, $p<.001$.

The final clustering score that was calculated was the D_A index proposed by Dalrymple-Alford (1970). In this case the O-E(SCR) deviation is expressed relative to the maximum possible

Table 3-10

Observed Clustering as a Proportion of Maximum Possible

SES Level	List No.	Trials 1	2	3	4	5	6	Means for: List No.	SES Level
High	1	.51	.53	.71	.73	.56	.63	.61	.58
	2	.38	.57	.68	.59	.59	.44	.54	
	1 & 2	.44	.55	.69	.66	.58	.54	.58	
Low	1	.50	.43	.70	.67	.63	.57	.58	.62
	2	.73	.55	.61	.74	.63	.61	.64	
	1 & 2	.62	.49	.65	.71	.63	.59	.62	

Table 3-11

Mean Adjusted Ratio of Clustering (ARC) Scores for the Categorized Lists

SES Level	List No.	Trials						Means for:	
		1	2	3	4	5	6	List No.	SES Level
High	1	.04	.22	.54	.59	.36	.48	.37	.32
	2	-.16	.33	.50	.55	.23	.13	.26	
	1 & 2	-.06	.27	.52	.57	.30	.30	.32	
Low	1	.00	-.03	.38	.42	.38	.33	.25	.37
	2	.54	.33	.41	.63	.48	.45	.47	
	1 & 2	.29	.16	.39	.53	.43	.39	.37	

range of SCR values (i.e., Max. SCR - Min. SCR) for a given S's recall sequence. The same three Ss that were excluded from the previous two analyses were dropped for the same reasons in this instance. Mean D_A index scores are given in Table 3-12. The analysis of variance scores as reported in Appendix W show exactly the same pattern as observed with the previous two ratio measures. The scores varied significantly with trials, $F(5,145) = 5.94$, $p < .005$, first increasing and then decreasing at the end.

The recall sequences were also analyzed to determine the extent to which subjective organization, or intertrial organization as it is also called, was imposed upon them by these Ss. This type of organization refers to the tendency of Ss to recall any two items together in contiguous positions in the recall sequence on two successive trials. Thus it is simply the repeated contiguity which defines a unit of organization. The basis for forming a unit does not have to be some cultural category built into the list by E. In fact, the basis of the unit may be entirely idiosyncratic to a particular S and completely unspecifiable by E. Subjective organization, however, can be, and was in the present study, scored for both C and NC lists.

Following Bousfield and Bousfield (1966) the unit of measurement of subjective organization is called the intertrial repetition (ITR). In the original work by Tulving (1962) and by Bousfield, Puff, and Cowan (1964) an observed unit of ITR, O(ITR), was scored only if two words appeared, not only contiguously, but in exactly the same order on two successive trials. This method is now referred to as scoring only unidirectional O(ITR) units, $O(ITR_U)$. Some subsequent work (e.g., see Mandler & Dean, 1969) has broadened the definition of the ITR measure to require only repeated contiguity of two items on the two trials. Thus, this method scores what are referred to as bidirectional units of observed ITR, $O(ITR_B)$. Since there has not as yet been any comparative work done with the unidirectional and bidirectional measures, both of them were used in the present study.

The raw data once again were the complete recall sequences including all types of errors. And, as is true with the clustering measures, O(ITR) scores of either type are not readily interpretable so deviations from expected values and ratio scores had to be used. Brief definitional formulas for all ITR scores are shown in Appendix X.

By far, the most typical procedure for evaluating O(ITR) scores has been to compare them with those values to be expected on the basis of chance, E(ITR) values. The score for comparisons between groups is the O-E(ITR) difference-score.

The E(ITR) values were first found by the formula provided by Bousfield and Bousfield (1966) in which the E(ITR) values are a function of the number of words recalled on each trial and the

Table 3-12

Mean D_A Index of Clustering Scores for the Categorized Lists

SES Level	List No.	Trials						Means for:	
		1	2	3	4	5	6	List No.	SES Level
High	1	.08	.14	.36	.40	.25	.33	.26	.23
	2	-.03	.22	.32	.37	.18	.10	.19	
	1 & 2	.03	.18	.34	.38	.22	.21	.23	
Low	1	.01	-.03	.26	.27	.26	.24	.17	.24
	2	.29	.22	.28	.42	.33	.32	.31	
	1 & 2	.16	.10	.27	.35	.29	.28	.24	

number of items common to recall on both trials. In the course of the work with these scores it was discovered that duplication errors had a substantial effect on these E(ITR) values. In contrast to the E(SCR) formula, the E(ITR) formula cannot, at the present anyway, accurately incorporate the occurrence of such errors. In the absence of analytic modifications of the formula, the mean of a series of 10 random permutations was used as a "corrected" E(ITR) value, and referred to here as E(ITR). More specifically, the items produced (including intrusions and duplications) by a S on a given pair of trials were randomly permuted 10 times. Each of the pairs of permutations were scored for observed ITR units. The number of units observed in these permutations was then taken as the estimate of the appropriate E values under these circumstances. It was found that, as would be expected, the E and E' values were quite similar when there were no duplication errors, and that as the number of duplication errors increased so did the value of E'(ITR) relative to E(ITR). In short, as the number of duplication errors increases, the value of E(ITR) progressively underestimates the value of E(ITR). Consequently, all of the analyses involving expected ITR values were done using both E and E' values. This, then, gives four basic difference-scores that were examined: O-E(ITR_U), O-E'(ITR_U), O-E(ITR_B), and O-E'(ITR_B). The mean scores for these measures are shown in Tables 3-13, 3-14, 3-15, and 3-16, respectively.

The degree of subjective organization imposed within each experimental condition was assessed by comparing the observed and expected amounts for each trial pair. The comparisons involved t-tests for correlated means and were performed separately for each of the four measures. The results of these comparisons are presented in Tables 3-17, 3-18, 3-19, and 3-20. There is much data in these tables, but since comparisons between conditions are made later, the important concern here is with the degree of organization within each condition. Perhaps the results can be summarized as follows. There appears to be some reduction in the number of significant deviations from chance when the corrected (E') chance values are used, especially in the later trials when the number of errors becomes quite high. Accordingly, it would seem best to emphasize the results based upon these corrected values (i.e., Tables 3-18 and 3-20). Considering the number and consistency of significant comparisons it would appear that a significant degree of subjective organization was imposed upon their output by both HI-SES and LO-SES groups in recall from C lists when considering either the unidirectionally ordered or bidirectionally unordered measures. However, with the NC lists the evidence for a significant usage of subjective organization is really only convincing in the case of the bidirectional measure. Overall then, there is at least some evidence that these Ss used subjective organization as one viable form of organization in recall.

Direct comparisons between the conditions of interest in this study were afforded by submitting each of the four sets of difference scores to an analysis of variance. The design for

Table 3-13

Mean O-E(ITR$_U$) Difference Scores

SES Level	Type of List	List No.	Trial Pairs 1-2	2-3	3-4	4-5	5-6	Means for: List No.	Type of List	SES Level
High	Categorized	1	.38	.39	.72	.74	1.73	.79	.77	.65
		2	.02	.57	.52	1.64	.95	.74		
		1 & 2	.20	.48	.62	1.19	1.34	.77		
	Non-Categorized	1	.27	-.33	.58	1.44	.96	.58	.53	
		2	.42	.06	1.05	.47	.35	.47		
		1 & 2	.34	-.14	.81	.96	.66	.53		
Low	Categorized	1	.05	.13	.95	1.10	2.11	.87	.89	.64
		2	.22	.56	.90	1.65	1.29	.92		
		1 & 2	.13	.35	.92	1.37	1.70	.89		
	Non-Categorized	1	.30	.34	.78	.33	.45	.44	.39	
		2	.09	.12	.54	.51	.41	.33		
		1 & 2	.19	.23	.66	.42	.43	.39		

Table 3-14

Mean O-E'(ITR_U) Difference Scores

SES Level	Type of List	List No.	Trial Pairs 1-2	2-3	3-4	4-5	5-6	Means for: List No.	Type of List	SES Level
High	Categorized	1	.41	.41	.62	.72	1.56	.74	.70	.50
		2	-.10	.56	.46	1.47	.88	.65		
		1 & 2	.16	.48	.54	1.09	1.22	.70		
	Non-Categorized	1	.31	-.46	.22	.59	.26	.18	.30	
		2	.42	.00	.94	.33	.39	.42		
		1 & 2	.37	-.23	.58	.46	.32	.30		
Low	Categorized	1	.01	.03	.66	.98	1.57	.65	.68	.45
		2	.14	.30	.80	1.49	.81	.71		
		1 & 2	.08	.17	.73	1.23	1.19	.68		
	Non-Categorized	1	.26	.19	.53	-.08	-.20	.14	.23	
		2	.11	.10	.51	.34	.49	.31		
		1 & 2	.18	.14	.52	.13	.14	.23		

Table 3-15

Mean O-E(ITR_B) Difference Scores

SES Level	Type of List	List No.	Trial Pairs 1-2	2-3	3-4	4-5	5-6	Means for: List No.	Type of List	SES Level
High	Categorized	1	.65	1.00	1.00	.58	2.24	1.09	1.01	.87
		2	.15	.80	.49	1.40	1.80	.93		
		1 & 2	.40	.90	.75	.99	2.02	1.01		
	Non-Categorized	1	-.02	-.22	1.04	2.11	1.59	.90	.73	
		2	.18	.35	.98	.94	.38	.57		
		1 & 2	.08	.06	1.01	1.52	.98	.73		
Low	Categorized	1	.32	.71	1.79	1.52	2.33	1.33	1.35	1.02
		2	.55	1.01	1.36	1.85	2.02	1.36		
		1 & 2	.43	.86	1.57	1.69	2.18	1.35		
	Non-Categorized	1	.48	.91	.90	1.22	1.00	.90	.70	
		2	.06	.13	.74	1.14	.38	.49		
		1 & 2	.27	.52	.82	1.18	.69	.70		

Table 3-16

Mean O-E'(ITR_B) Difference Scores

SES Level	Type of List	List No.	Trial Pairs 1-2	2-3	3-4	4-5	5-6	Means for: List No.	Type of List	SES Level
High	Categorized	1	.66	1.13	.74	.69	1.84	1.01	.92	.66
		2	.14	.70	.46	1.24	1.61	.83		
		1 & 2	.40	.92	.60	.97	1.73	.92		
	Non-Categorized	1	-.02	-.41	.60	.64	.60	.28	.39	
		2	.32	.20	.83	.72	.41	.50		
		1 & 2	.15	-.11	.72	.68	.51	.39		
Low	Categorized	1	.20	.62	1.24	1.08	1.56	.94	.93	.64
		2	.48	.39	1.12	1.47	1.11	.91		
		1 & 2	.34	.51	1.18	1.27	1.33	.93		
	Non-Categorized	1	.37	.64	.33	.26	-.03	.31	.35	
		2	.03	.03	.70	.76	.43	.39		
		1 & 2	.20	.34	.52	.51	.20	.35		

Table 3-17

Results of t-tests of O-E(ITR_U) Differences

SES Level	Type of List	List No.	Results	Trial Pairs 1-2	2-3	3-4	4-5	5-6	df
High	Categorized	1	t	1.69	.75	2.27	2.31	3.47	8
			p			†	*	**	
		2	t	.16	1.96	1.43	2.70	1.96	8
			p		†		*	†	
		1 & 2	t	1.51	1.66	2.64	3.39	3.83	17
			p			*	**	**	
	Non-Categorized	1	t	.71	1.87	1.37	2.18	2.47	8
			p		†		†	*	
		2	t	1.86	.36	1.79	1.30	1.10	8
			p	†					
		1 & 2	t	1.62	-1.03	2.29	2.62	2.56	17
			p						
Low	Categorized	1	t	.20	.38	1.36	4.43	3.26	8
			p				**	*	
		2	t	1.21	2.97	2.94	1.98	3.71	8
			p		*	*	†	**	
		1 & 2	t	.92	1.75	2.50	3.21	4.60	17
			p		†	*	**	***	
	Non-Categorized	1	t	1.21	1.31	3.94	1.42	1.19	8
			p			**			
		2	t	.40	.70	1.68	2.23	1.24	8
			p				†		
		1 & 2	t	1.20	1.50	3.57	2.63	1.77	17
			p						

† $.10 > p > .05$
* $p < .05$
** $p < .01$
*** $p < .001$

Table 3-18

Results of t-tests of C-E'(ITR_U) Differences

SES Level	Type of List	List No.	Results	Trial Pairs 1-2	2-3	3-4	4-5	5-5	df
High	Categorized	1	t	1.61	.91	2.18	2.03	3.13	
			p			†	†	*	8
		2	t	.09	1.64	1.70	2.51	1.50	
			p				*		8
		1 & 2	t	1.40	1.64	2.81	3.12	3.23	
			p			*	**	**	17
	Non-Categorized	1	t	.40	-1.98	1.07	1.59	.32	
			p		†				8
		2	t	1.72	.50	1.65	1.19	1.19	
			p						8
		1 & 2	t	1.45	-.95	1.97	2.02	1.52	
			p			†	†		17
Low	Categorized	1	t	-.39	.50	1.48	3.81	2.82	
			p				**	**	8
		2	t	1.61	1.40	2.63	2.04	2.84	
			p			*	†	*	8
		1 & 2	t	.77	1.29	2.85	2.94	3.57	
			p			*	**	**	17
	Non-Categorized	1	t	.83	1.58	3.49	-.54	-.56	
			p			**			8
		2	t	.30	.11	1.49	1.72	1.39	
			p						8
		1 & 2	t	.86	1.29	2.67	.48	.51	
			p			*			17

† .10>p>.05
* p<.05
** p<.01
*** p<.001

Table 3-19

Results of t-tests of O-E(ITR_B) Differences

SES Level	Type of List	List No.	Results	Trial Pairs 1-2	2-3	3-4	4-5	5-6	df
High	Categorized	1	t	2.01	2.31	3.57	1.84	4.46	8
			p	†	*	**		**	
		2	t	.64	2.98	1.09	2.38	3.09	8
			p		*		*	*	
		1 & 2	t	1.97	3.63	2.83	2.93	5.36	17
			p	†	**	*	**	***	
	Non-Categorized	1	t	-.08	-.64	2.86	3.22	2.34	8
			p			*	*	*	
		2	t	.97	2.31	2.02	2.40	1.20	8
			p		*	†	*		
		1 & 2	t	.44	.32	3.43	3.84	2.51	17
			p			**	**	*	
Low	Categorized	1	t	1.41	2.10	2.08	5.07	4.46	8
			p		†	†	***	**	
		2	t	3.25	2.52	3.21	2.42	3.69	8
			p	*	*	*	*	**	
		1 & 2	t	3.12	3.34	3.36	4.21	5.90	17
			p	**	**	**	***	***	
	Non-Categorized	1	t	1.49	2.24	3.41	2.92	1.89	8
			p		†	**	*	†	
		2	t	.25	.56	3.33	7.91	1.08	8
			p			*	***		
		1 & 2	t	1.33	2.11	4.86	5.50	2.18	17
			p		*	***	***	*	

† .10>p>.05
* p<.05
** p<.01
*** p<.001

Table 3-20

Results of t-tests of O-E'(ITR_B) Differences

SES Level	Type of List	List No.	Results	Trial Pairs 1-2	2-3	3-4	4-5	5-6	df
High	Categorized	1	t	1.92	2.40	2.77	1.69	4.44	8
			p	†	*	*		**	
		2	t	.18	2.47	1.16	2.04	2.44	8
			p		*		†	*	
		1 & 2	t	1.54	3.34	2.66	2.65	4.57	17
			p		**	*	*	***	
	Non-Categorized	1	t	-.94	-1.79	2.01	1.88	1.73	8
			p			†	†		
		2	t	.78	1.23	1.78	2.33	1.17	8
			p				*		
		1 & 2	t	-.21	-.66	2.73	3.04	2.11	17
			p			*	**	*	
Low	Categorized	1	t	.89	1.85	3.34	3.34	3.44	8
			p			*	*	**	
		2	t	2.94	1.72	2.56	2.25	2.55	8
			p	*		†	†	*	
		1 & 2	t	2.40	2.60	4.25	3.38	4.39	17
			p	*	*	***	**	***	
	Non-Categorized	1	t	1.37	2.74	2.42	.87	-.13	8
			p		*	*			
		2	t	.22	.12	3.20	7.50	1.20	8
			p			*	***		
		1 & 2	t	1.16	2.02	4.00	2.93	.76	17
			p		†	***	**		

† $.10 > p > .05$
* $p < .05$
** $p < .01$
*** $p < .001$

these analyses includes SES level, type of list, and list-pair no. as between-Ss variables, while 5 trial-pairs comprised the within-Ss variable. Detailed results of the analyses of O-E(ITR_U), O-E'(ITR_U), O-E(ITR_B), and O-E'(ITR_B) are shown in Appendices Y, Z, AA and BB, respectively.

All four of these analyses showed exactly the same pattern of results. The only significant effects were type of list, trial-pairs, and the interaction of these two variables. Thus, the amount of observed subjective organization exeeeded the chance amount to a significantly greater extent in the C than the NC lists, and while there was a significant overall increase in the deviation over trial-pairs, the rate of increase was more rapid with the C than the NC lists.

A score representing the observed amount of subjective organization as a proportion of the maximum possible amount was also determined in the manner shown in Appendix X. Once again, scores were found for both unidirectionally ordered units and for bidirectional units. However, since chance expectations do not enter into the calculation of these scores there did not, of course, have to be separate results for E and E'. Mean values of the Ratio ITR_U scores are shown in Table 3-21, and those for Ratio ITR_B are presented in Table 3-22. It can be seen that the overall average degree of observed subjective organization within conditions begins at a very low level relative to the maximum possible, and while it increases over trial-pairs, it does not exceed a level of about 21% in the case of unidirectional units or about 35% in the case of bidirectional units.

Comparisons between conditions were afforded by means of the analysis of variance. The detailed results of these analyses are presented in Appendix CC for Ratio ITR_U scores and in Appendix DD for Ratio ITR_B scores. Both analyses revealed significant increases in scores over trial-pairs and significant ($p<.05$) four-way interactions. These were the only significant effects in the analysis of the Ratio ITR_U scores. However, the analysis of the Ratio ITR_B scores also indicated a significant type of list x trial pairs interaction, F (4,256) = 3.03, $p<.025$, which by inspection seems to reflect several reversals of which type of list is superior across trial-pairs. In addition, the SES x type of list x list-pairs, and the SES x type of list x trial-pairs interactions were marginally ($.10>p>.05$) significant. These marginal effects were not further tested, but seemed to reflect same trend toward greater Ratio ITR_B scores for LO-SES Ss on the NC2 list and for HI-SES Ss on the NC1 list.

A subjective organization score conceptually analogous to the adjusted ratio of clustering score was found, as shown in Appendix X, by taking the observed deviation from the chance expected values relative to the maximum possible deviation from chance for a given recall sequence. Separate scores were found for unidirectional and bidirectional units, and since

Table 3-21

Observed ITR_U Scores as a Proportion of Maximum Possible Values

SES Level	Type of List	List No.	Trial Pairs 1-2	2-3	3-4	4-5	5-6	Means for: List No.	Means for: Type of List	Means for: SES Level
High	Categorized	1	.16	.15	.23	.18	.27		.20	.20
		2	.08	.21	.19	.35	.17	.20		
		1 & 2	.12	.18	.21	.27	.22	.20		
	Non-Categorized	1	.13	.03	.21	.39	.28	.21	.19	
		2	.21	.12	.24	.17	.15	.18		
		1 & 2	.17	.07	.23	.28	.22	.19		
Low	Categorized	1	.06	.13	.15	.25	.28	.18	.19	.20
		2	.15	.14	.22	.24	.24	.20		
		1 & 2	.10	.14	.18	.25	.26	.19		
	Non-Categorized	1	.15	.16	.24	.13	.13	.16	.21	
		2	.17	.27	.30	.38	.19	.26		
		1 & 2	.16	.21	.27	.26	.16	.21		

Table 3-22

Observed ITR_B Scores as a Proportion of Maximum Possible Values

SES Level	Type of List	List No.	Trial Pairs 1-2	2-3	3-4	4-5	5-6	Means for: List No.	Type of List	SES Level
High	Categorized	1	.30	.38	.37	.22	.43	.34	.34	.34
		2	.21	.53	.26	.38	.35	.35		
		1 & 2	.26	.45	.31	.30	.39	.34		
	Non-Categorized	1	.15	.19	.47	.63	.44	.37	.34	
		2	.25	.34	.33	.37	.23	.30		
		1 & 2	.20	.26	.40	.50	.33	.34		
Low	Categorized	1	.22	.29	.46	.37	.40	.35	.34	.35
		2	.29	.28	.36	.34	.36	.32		
		1 & 2	.26	.28	.41	.35	.38	.34		
	Non-Categorized	1	.28	.36	.33	.33	.28	.31	.37	
		2	.39	.37	.38	.70	.31	.43		
		1 & 2	.34	.36	.35	.52	.30	.37		

expected values enter into these calculations separate scores also had to be found for E and E' units. Thus, there were four types of adjusted ratio of subjective organization (ARSO) scores: $ARSO_U$, $ARSO'_U$, $ARSO_B$, and $ARSO'_B$. Mean values of each of these types of scores are given in Tables 3-23, 3-24, 3-25, and 3-26.

Analyses of variance were performed upon each set of scores with results as described in Appendices EE, FF, GG, and HH. Once again, all four analyses showed basically the same pattern of results. There was no significant main effect of SES, nor did SES enter into any significant interactions except for the four-way interaction of all variables in the case of the $ARSO_U$ and $ARSO'_B$ scores. This four-way interaction was marginally significant ($.10>p>.05$) with the $ARSO'_U$ scores, and dropped below even this level with the $ARSO_B$ scores. Furthermore, as was found with all of the O-E difference scores and the Ratio ITR_B score, the $ARSO_B$ score results indicated a significant interaction between type of list and trial-pairs. This appeared to be due to the fact that the C lists were superior on the first two trial-pairs, inferior to the NC lists on the middle two trial-pairs, and superior once again on the last trial-pair. Finally, the analysis of these $ARSO_B$ scores showed marginally significant type of list x list-pair and SES x type of list x list-pair interactions--suggesting a trend toward higher scores for the LO-SES Ss on NC2 and C1 but lower scores on NC1 and C2.

Discussion

Because of the large number of analyses and the complexity of certain of the results, the basic findings will be summarized in some detail.

Firstly, there were no significant overall differences in the number of stimulus words recalled correctly by the two SES groups. However, there was a significant SES x type of list x list-pair no. interaction such that more words were recalled correctly by the HI-SES Ss from the NC2 list whereas the LO-SES Ss recalled more from the C2 list. Furthermore, a tendency toward greater correct recall from both C lists by the LO-SES groups and from both NC lists by the HI-SES groups was suggested by the marginally significant SES x type of list interaction. Looked at in terms of the C vs. NC list differences, the interactions indicate that more words were recalled from C2 than NC2 by the LO-SES Ss, but not the HI-SES Ss, and the C vs. NC discrepancy tended to be greater in both list pairs for the LO-SES Ss.

The error data revealed that the overall number of all kinds of errors combined reached a level of between 17 and 35% of the total responses produced. The majority of the errors were duplications of previously emitted responses rather than intrusions of items which were not in the stimulus list. The analysis of the log-transformed error data indicated that the LO-SES groups made significantly more errors overall than did the HI-SES groups. The significant SES x type of list x list pair no. interaction

Table 3-23

Mean Adjusted Ratio of Subjective Organization ($ARSO_U$) Scores

SES Level	Type of List	List No.	Trial Pairs 1-2	2-3	3-4	4-5	5-6	Means for: List No.	Type of List	SES Level
High	Categorized	1	.08	.05	.17	.10	.19	.12	.12	.11
		2	-.03	.11	.11	.29	.09	.11		
		1 & 2	.03	.08	.14	.19	.14	.12		
	Non-Categorized	1	.00	-.10	.12	.31	.19	.11	.10	
		2	.13	.03	17	.09	.08	.10		
		1 & 2	.07	-.04	.14	.20	.14	.10		
Low	Categorized	1	-.02	.05	.08	.20	.23	.11	.12	.12
		2	.06	.07	.15	.17	.18	.13		
		1 & 2	.02	.06	.11	.19	.20	.12		
	Non-Categorized	1	.03	.06	.17	.04	.06	.07	.12	
		2	.03	.17	.26	.31	.08	.17		
		1 & 2	.03	.12	.22	.18	.07	.12		

Table 3-24

Mean Adjusted Ratio of Subjective Organization (ARSO'$_U$) Scores

SES Level	Type of List	List No.	Trial Pairs 1-2	2-3	3-4	4-5	5-6	Means for: List No.	Type of List	SES Level
High	Categorized	1	.10	.05	.14	.11	.19	.12	.11	.09
		2	-.11	.15	.09	.27	.07	.09		
		1 & 2	-.00	.10	.12	.19	.13	.11		
	Non-Categorized	1	.04	-.13	.04	.24	.13	.07	.08	
		2	.13	.01	.15	.06	.09	.09		
		1 & 2	.09	-.06	.09	.15	.11	.08		
Low	Categorized	1	-.04	-.02	.04	.20	.18	.07	.08	.09
		2	.01	.04	.14	.16	.12	.10		
		1 & 2	-.01	.01	.09	.18	.15	.08		
	Non-Categorized	1	.02	.02	.12	-.01	-.03	.02	.09	
		2	-.05	.18	.27	.24	.11	.15		
		1 & 2	-.01	.10	.19	.12	.04	.09		

Table 3-25

Mean Adjusted Ratio of Subjective Organization ($ARSO_B$) Scores

SES Level	Type of List	List No.	Trial Pairs 1-2	2-3	3-4	4-5	5-6	Means for: List No.	Type of List	SES Level
High	Categorized	1	.23	.22	.37	.04	.29	.24	.24	.22
		2	.01	.42	.20	.24	.34	.24		
		1 & 2	.14	.32	.29	.14	.31	.24		
	Non-Categorized	1	-.05	-.06	.33	.54	.27	.21	.20	
		2	.06	.19	.18	.35	.20	.19		
		1 & 2	.01	.07	.25	.44	.24	.20		
Low	Categorized	1	.29	.12	.36	.27	.29	.27	.23	.24
		2	.14	.14	.24	.21	.25	.20		
		1 & 2	.22	.13	.30	.24	.27	.23		
	Non-Categorized	1	.06	.18	.20	.17	.15	.15	.25	
		2	.23	.18	.65	.62	.10	.36		
		1 & 2	.15	.18	.42	.40	.12	.25		

Table 3-26

Mean Adjusted Ratio of Subjective Organization ($ARSO'_B$) Scores

SES Level	Type of List	List No.	Trial Pairs 1-2	2-3	3-4	4-5	5-6	Means for List No.	Means for Type of List	Means for SES Level
High	Categorized	1	.17	.25	.22	.07	.27	.20	.19	.16
		2	-.00	.38	.09	.24	.19	.18		
		1 & 2	.08	.31	.15	.16	.23	.19		
	Non-Categorized	1	-.09	-.07	.26	.38	.18	.13	.13	
		2	.11	.12	.10	.21	.09	.13		
		1 & 2	.01	.03	.18	.30	.13	.13		
Low	Categorized	1	.03	.07	.33	.22	.21	.17	.15	.14
		2	.07	.06	.21	.17	.16	.14		
		1 & 2	.05	.07	.27	.20	.19	.15		
	Non-Categorized	1	.03	.11	.06	.03	.01	.05	.13	
		2	-.12	.19	.31	.55	.16	.21		
		1 & 2	-.05	.15	.19	.29	.08	.13		

appears to reflect the fact that the largest SES discrepancy was in C list recall, and especially in the case of C1, with nearly equal acores on NC2. This pattern seems most pronounced in the middle trials since there was a big jump in errors for the HI-SES groups on the last trial (i.e., the SES x type of list x list pair no. x trials interaction was significant).

Insignificant overall differences between the two SES levels were also obtained in the analysis of the total number of responses produced in recall. However, after the first trial, the LO-SES Ss emitted a consistently higher mean number of total responses, though the significant SES x trials interaction and post hoc comparisons revealed that the difference was significant only on the fifth trial. The marginally significant SES x type of list and SES x type of list x list-pair no. interactions suggested some trend toward more items being emitted by LO-SES groups in recall from C lists, but not NC lists, and this trend was somewhat greater with the second pair of lists.

As suggested by the pattern of previous results, the analysis of the proportion of the total number of items emitted which were actually correct revealed significantly higher overall scores for the HI-SES groups. These scores showed a significant decrease over trials (corresponding to the increase in errors) and a significant four-way interaction which does not appear to be readily describable.

The evidence concerning the occurrence of categorical clustering in recall of the C lists indicated substantial use of common conceptual relationships for organizing reaall by all Ss. That is, after the first two trials, nearly all mean observed clustering scores deviated significantly from chance expectations. Furthermore, observed clustering as a proportion of the maximum possible amount began at an overall average of about 53% and attained a highest overall amount of about 69% on the fourth trial, but showed somewhat of a decline on the last two trials where the number of errors reached such a high level. None of the comparisons between groups was significant regardless of whether the O-E(SCR) deviation scores, the Ratio SCR scores, the ARC scores, or the D_A Index scores were examined. The similarity of results with all of these different types of scores thus affords a degree of confidence that the conclusions are not dependent upon the statistical properties of any particular index of clustering.

The pattern of usage of subjective organization was somewhat more complex for several reasons. The extent to which observed subjective organization (in ITR units) deviated from chance expectations supported the conclusion that a significant amount of this kind of organization was imposed in quite a few cases, especially on the last three trial-pairs. While scores based on the standard formula for chance, E(ITR), units were analyzed, it is proposed that the scores based upon chance values estimated from randomized sequences, E'(ITR) units, are more appropriate in an instance like the present one where there is such a high number of errors. Very few of these deviations were significant in the first two trial-pairs. The results for the last three trial-pairs were somewhat more impressive when counting bidirectional units rather than just unidirectional, and for the C as opposed to the NC lists (i.e., the highest ratio of significant comparisons was 13/18 with the O-E'(ITR_B) scores for the C lists while the poorest ratio was 2/18 in the case of the O-E'(ITR_U) scores for the NC lists). The analyses of these four sets of deviation scores all confirmed that the only significant effects were associated with the increase over trial-pairs and the superiority of scores for the C lists.

The Ratio ITR scores expressed the observed amount of subjective organization relative to the maximum possible amount. Averaged over everything these scores showed observed organization to begin at about 21% of the maximum possible, and to reach a peak level of about 36%, followed by somewhat of a decline. Scores for the bidirectional measures were approximately twice as large as those for unidirectional measures, and in both cases the difference between the C and NC lists disappeared. The only evidence of SES effects was an undescribable significant four-way interaction with the Ratio ITR_U scores and marginally significant SES x type of list x trial-pairs and SES x type of list x list pair no. interactions in the case of Ratio ITR_B scores.

The adjusted ratio of subjective organization (ARSO) scores showed much the same pattern of results. Again, the only evidence of SES differences was given by significant four-way interactions with the $ARSO_U$ and $ARSO'_B$ scores and a marginally significant one with the $ARSO'_U$ scores. The $ARSO_B$ scores showed a marginally significant SES x type of list x list-pair no. interaction which seemed to indicate some trend toward higher scores for LO-SES Ss on NC2 and C1 and for HI-SES Ss on NC1 and C2.

Thus, the various subjective organization measures, like those for clustering, were in agreement in providing only the most minimal indication of any SES effects (the four-way interactions). And the agreement among the measures again provides a degree of confidence of freedom from statistical artifacts.

These results appear to bear upon several aspects of Jensen's analysis. Jensen's work led to the expectation that the HI-SES and LO-SES groups should have been equivalent in both recall and organization in both the C and NC types of lists, though for somewhat different reasons. Equal performance on the NC lists is predicted on the grounds that he has theorized that mastery of these lists requires only Level I abilities and they are proposed to be equally distributed within all SES levels. The mastery of C lists, on theother hand, is thought to require the more complex Level II abilities which are distributed about a higher mean in higher SES groups. However, on the basis of his previous data, Jensen has proposed that the superior Level II abilities of the HI-SES Ss only develop sometime after the kindergarten level (but before the fifth grade). Thus, there should have been no differences in either Level I or Level II abilities between the SES groups at the nursery school level studied here.

In contrast to these expectations, the present results indicate some substantial differences in the recall performance of the two SES groups. Even the previously used measure, the number of words recalled correctly, showed such differences. Thus, the HI-SES Ss recalled significantly more words correctly from the NC2 list when they should have been quite equivalent to the LO-SES Ss in the Level I abilities required to master such a list. An even clearer contradiction is provided by the finding that the LO-SES Ss recalled significantly more words correctly from the C2 list when they should have been, at the very best, only equal to the HI-SES Ss in the theorized important Level II abilities at this stage of development. Furthermore, there was a marginally significant trend toward greater recall from both C lists by the LO-SES groups and from both NC lists by the HI-SES groups. The interaction with the precise set of stimulus materials of both the C and NC types is bothersome, since if only the results of the C1 and NC1 lists are considered the findings do conform to Jensen's predictions. However, Jensen's theorizing does not include any qualifications about properties of stimulus materials other than their basic C or NC structure.

The other measures of recall performance also show unpredicted differences between SES groups which should be equivalent in terms of both of Jensen's postulated kinds of relevant abilities. Thus, the LO-SES groups committed significantly more errors overall, and emitted more total responses on the fifth trial. On the other hand, the HI-SES Ss had significantly higher overall proportions of their recalled items which were correct.

The data on the two forms of organization in recall are also relevant to Jensen's theorizing in several ways. In the first place, Jensen has proposed that the occurrence of category clustering in the recall of C lists represents a major transformation of the input order in accordance with learned concepts, and thus reflects the operation of the Level II abilities. As noted before, Level II abilities are thought to distinguish between SES groups, but not to develop even in HI-SES Ss until some time after the kindergarten age. The present findings of no significant differences in the amount of category clustering between the two SES groups are therefore consistent with Jensen's position. However, the clustering data also revealed that the lack of difference between the SES groups arose because both groups imposed equally significant amounts of clustering, rather than because neither group imposed any significant clustering. The significant deviations of observed from expected scores demonstrate that the ability for clustering is already present in these nursery school children, and to an equal extent for both SES groups. Furthermore, the comparison of observed clustering as a percentage of the maximum possible amount showed the present Ss to begin at a level of about 53% whereas the average over three groups of college students in a study by Puff (1970) was just about 54%. The college students, however, did increase to a level of about 79% after six trials while the present Ss had attained the somewhat lower peak level of 69% and dropped off somewhat on the next two trials. The present data, then, strongly suggest that while the level of clustering may increase significantly with age, even these nursery school Ss have demonstrated a significant clustering tendency. The implication of these findings appears to be that either category clustering is not a reflection of Jensen's Level II ability, or this ability appear earlier than he has thought and is equally present in both SES populations.

The present data on subjective organization also seem inconsistent with some of Jensen's analysis. Jensen has not made any explicit statements about this form of organization, but several kinds of expectations presumably follow from what he has said. He has suggested that in the recall of C lists LO-SES Ss tend to produce idiosyncratic units of organization based upon functional associations, rather than the type of conceptual clustering produced by HI-SES Ss. It is assumed here that the subjective organization measure should be an appropriate way to take into account the use of reliable idiosyncratic units. Furthermore, it is deduced from what Jensen has said that this form of organization is not to be regarded as reflecting Level II abilities. It should thus appear developmentally earlier than conceptual clustering, and should be equally distributed across SES groups. Consistent with this, the analyses of varieus subjective organization measures revealed only the most minimal SES differences (i.e., only in four-way interactions) in either type of stimulus list. Once again, however, the degree of organization within each condition is of particular interest. The deviation of observed from expected ITR scores indicated that a significant amount of this type of organization was imposed in later trials according to the outcome of a fair number of the individual comparisons. In comparison with the results for clustering, however, the evidence for the degree of occurrence of subjective organization is not nearly as impressive, especially with the NC lists. This suggests the possibility that clustering which is postulated to reflect Level II abilities, and to appear developmentally later, occurred to a greater extent than subjective organization which presumably does not reflect anything more than Level I abilities and should appear earlier.

Several other findings, however, tend to complicate this interpretation. For one thing, the subjective organization measures, in the case of the C lists, may be tapping more units of words from the same category rather than the kind of idiosyncratic functional associations mentioned by Jensen. No analysis was directed at this question, but it was found that subjective organization as specified by the O-E measure was significantly greater for the C than the NC lists. Secondly, the comparison of the observed unidirectional units relative to the maximum possible organization shows these Ss to begin at virtually the same level (11%) as Puff's college students and then to reach a peak of about 27% in 5 trial-pairs whereas the college Ss reach only about 16% in the same number of trials. None of the subject populations, then, relied very heavily upon this form of organization.

In addition to the foregoing empirical analysis of subjective organization, it seems necessary to add that it is difficult to conceive how this form of organization really could be thought to involve a lower level ability than does clustering. If the significant transformation of the input by the S is the critical defining characteristic of clustering as a Level II ability, then subjective organization must also represent a reflection of Level II ability for it also involves a substantial transformation of input, and a reliable transformation which occurs on two successive trials. As indicated before, Jensen has not yet explicitly considered this form of organization in his theoretical views of SES differences, but it would appear to be time that he did so.

Overall, then, the combined evidence concerning the two forms of organization provides little support for the view that the recall performance results can be explained on the basis of the organization of recall. That is, in terms of recall performance the HI-SES Ss recalled more words from NC2; the LO-SES Ss recalled more from C2; the LO-SES Ss committed more errors; and the HI-SES Ss had a higher proportion of recall which was correct. However, there were no corresponding significant variations in either form of organization. Thus, it would be difficult to argue that recall performance depended upon the kinds of mediational activities presumably indexed by either of the forms of organization measured here. These findings not only run counter to Jensen's position in the ways already indicated, but also offer little encouragement for generalizing to cultural deprivation effects from, for example, Spitz's (1966) view that the learning and memory deficit of the mentally retarded child results from a failure to organize adequately. Nor do these findings afford any generality for the general notion, developed mainly with college age Ss, that recall is always dependent upon organization (e.g., see Mandler, 1967; and Tulving, 1962, 1968). It is, of course, possible that there are still some critical mediational processes which are in fact indexed by some measurable form of organization of recall, but that this crucial form of organization has not been specified yet.

The findings of this study similarly do not seem to provide any support for any interpretation of LO-SES performance in terms of inhibition due to the "strangeness" of the situation; reduced motivation because of perceived "irrelevance" of the task; a lower level of verbal output; or a reduced level of spontaneous production of mediating responses of the kind thought to be important in organization of recall. The principal data which argue against these kinds of views is the finding that the LO-SES Ss produced as many total responses in the situation as did the HI-SES Ss, and even emitted a significantly greater total number of responses on the fifth trial. They were also quite equivalent in organization which presumably requires substantial effort and "processing" on the part of the Ss.

The pattern of recall results does, however, indicate an interesting difference between the performance of HI-SES and LO-SES groups. The two groups were essentially equivalent overall in the number of items recalled correctly, though the HI-SES Ss were superior with the NC2 list and the LO-SES Ss were superior with the C2 list. The LO-SES Ss did, however, produce more errors, causing them to also show a somewhat greater number of total responses produced. In contrast, the HI-SES Ss produced a higher proportion of their total items which were correct. This pattern suggests the interpretation that the HI-SES Ss are superior in their ability or inclination to edit their recall processes and to filter out erroneous responses. The breakdown of errors showed a preponderance of duplications over intrusions. Thus, it must be that the HI-SES Ss were superior in their ability to remember the responses previously emitted on a given trial and to avoid duplicating them later on that trial. Whether the difference reflected here is one of ability or motivation cannot, of course, be determined from the present data. It is possible that LO-SES Ss are deficient in the capability for this type of editing, selecting, or filtering, but it is also possible that their previous experience is somehow such that minimizing errors is just not one of the demand characteristics of the experimental situation as it is perceived by them. In either case, the critical SES-related determinants are quite uncertain.

Essentially this kind of a selector mechanism has been previously proposed (e.g., Underwood & Schulz, 1960), and results in support of such a mechanism were obtained with college students by Bousfield and Rosner (1970). These investigators compared oral recall of a NC list by a group givan standard (S) free recall instructions and a group given instructions for what they called uninhibited (U) recall. The Ss in the U condition were basically told to say any words that came into their minds even if they knew that they would be making errors. Their results showed that Ss in the U group did make significantly more errors and most of them were duplications. Both groups recalled about the same number of words correctly on the first trial, but on the fifth trial the U group recalled significantly more. Finer analyses traced the effect to the fact that the U Ss showed less intertrial forgetting. In other words, the Ss who filtered substantially also apparently mistakenly edited out some correct responses which had been given on the previous trial as well as avoiding duplication of responses given earlier in recall on the same trial. Thus, Bousfield and Rosner's results are basically similar in several respects to the present pattern of SES differences, thereby lending some support to the notion that some type of editing may have been involved here. Their findings further suggest that a high degree of editing is probably not an optimal strategy for maximizing recall scores in a situation where there is no penalty for making errors. They state (p. 76) that "these results imply that standard instructions may implicitly impose a strong inhibition of errors that generalized to potentially and previously accessible items. If this effect indeed obtains then studies which aim at determining MTFR capacity should provide instructions which, in effect, stipulate the permissibility of emitting intratrial item repetitions." The really interesting implication of these findings is that in the present experiment the LO-SES Ss may have used a more optimal strategy (albeit unintentionally) and their correct recall scores may closely approximate their maximal memory capacity. On the other hand, to the extent that they edited more, and that this is not an optimal strategy, the scores for the HI-SES Ss may underestimate their true capacity. That still leaves the question, of course, as to why the HI-SES Ss may be following a nonoptimal strategy.

In conclusion, the results of this study are not very supportive of Jensen's theorizing about SES differences in terms of Level I and Level II abilities, nor do they agree very well with some of the evidence previously provided by him and his colleagues. The reasons for this are certainly not clear, but several kinds of advantages can be claimed for the present investigation. Perhaps most importantly, confounding of race and SES was avoided here and has typically been present in Jensen's work. This study also involved a more detailed assessment of recall performance and a finer grained analysis of organization of recall then has previously been presented.

Several kinds of difficulties were also encountered here. Some quite unexpected and unexplainable differences between supposedly equivalent sets of stimulus materials were observed. Thus, for example, more words were recalled correctly by HI-SES Ss from the NC2 list but not the NC1 list, while the LO-SES Ss were superior with the C2 list, but not the C1 list. These findings suggest the need for future studies to continue to employ more than a single set of materials for protection, or else to investigate the role of stimulus materials directly. Perhaps the best procedure would be to collect normative data directly from the population to be involved in the experiment so that it would be certain that the materials were equally familiar, meaningful, interassociated, etc. for Ss at the different SES levels. Furthermore, to the extent that the present interpretation of a higher degree of editing, or filtering, by the HI-SES Ss is accurate, and if the interest in memory capacity difference remains, it would seem advantageous to include a condition in future experiments where the instructions encourage "uninhibited" recall. Beyond this, the origins of any such differences in filtering need to be investigated, and the operation of such a mechanism needs to be looked at developmentally. Finally, the finding that the standard expected values of subjective organization are substantially incorrect when there are large numbers of duplication errors will need to be taken into account in future studies with subject populations like these. Perhaps some analytical work can specify how the formula can be modified so as to avoid the necessity of estimating values from statistical subjects. Thus, there are a number of potentially fruitful directions for further research in this area.

Experiment 3: Cross Modal Equivalence in Nursery School Children from Different Socioeconomic Backgrounds

Purpose

Based upon casual observation that "culturally disadvantaged" children frequently appear considerably brighter than "equal IQ" middle class children, Jensen initiated a research program which culminated in his now famous article in the Harvard Educational Review (1969). That same year he published a more complete description of his two-level theory of learning to account for the social class difference in intelligence and learning (Jensen, 1969b).

Briefly, Jensen postulates the existence of two functionally related but genetically distinct mental processes which are basic to performance in various tasks. Level I abilities are basic to performance on tasks such as rote memory and serial learning; tasks which are relatively simple and are dependent upon rather low level associative learning. Level II abilities are fundamental to performance on tasks of greater complexity such as categorical free recall and Raven's progressive matrices; tasks which demand some subject initiated symbolic activity. Jensen conceives these abilities as genotypically distinct processes which are based upon different biological or structural foundations. The performance on various tasks is the phenotypical indication of the genotypic process underlying that performance.

Jensen also postulates that Level I and Level II processes are distributed around different means in children from different social classes. While Level I abilities are equally distributed in children from various social classes Level II abilities are distributed around a higher mean in children from upper socioeconomic classes relative to children from lower classes. A great deal of the research efforts of Jensen and others (Glassman, 1968; Rapier, 1968) have been oriented to the specification of tasks which are valid indicators of Level II abilities.

Jensen (1969a) speculates that perhaps the most pure and simple instance of the abilities necessary for performance on Level II tasks is cross modal transfer (CMT). CMT has been defined by Blank and Bridger (1964) as the ability to utilize information acquired through one sense modality in the solution of a problem presented in a second modality. These investigators distinguished between cross modal equivalence (CME) and cross modal concept (CMC). CME is the recognition of a specific stimulus in two different sensory modalities while CMC involves the application of a concept (e.g., oddity) in one modality after having acquired it in a different modality.

Millar (1971) presented nonsense shapes to 3 and 4 year old children through either the visual, haptic (active touch) or a combination of these modes of sensory input. She asked these Ss to select from an array of 4 forms presented to either visual, haptic or a combination of the two modes of sensory input, that stimulus which had just been seen/felt or seen and felt. She concluded that cross modal equivalences were formed by her children and that the visual modality was the more efficient modality. Birch and Lefford (1963) assessed the existence of cross modal equivalence in children employing the visual, haptic (active touch) and kinesthetic (hand movement tracing shape outline) sensory modes. They concluded that CME existed in children as young as 5 years of age and found CME ability to increase with age.

Connors, Schuette and Goldman (1967) investigated haptic to visual transfer in children ranging in age from kindergarten through sixth grade from both middle and lower class families. They concluded that such transfer ability is complete by age 6 for the dimension of shape but not for size or orientation. They also found a deficit in the lower class children along each of the three dimensions. It should be noted that children were randomly selected from two schools, one of which was located in "a slum area of the city which serves children who live in public low-cost housing developments or local dilapidated single dwellings." The other school was located "near the suburbs of Baltimore and serves children whose families reside for the most part in wealthy neighborhoods." In this study, as in several others (e.g., Glassman, 1968) social class was not measured, IQ was not controlled, and race or ethnic-background was confounded with SES.

The present study assesses both within modal and cross modal equivalence in nursery school children differing in the socio-economic status (SES) of their parents. Two groups of children, one from high SES families and the other from low SES families were formed. The children were English speaking, caucasian children who were assigned to groups which were equated for PPVT performance levels.

Each SES group was split into a tactual and visual subgroup. The tactual Ss experienced the first or standard stimulus in each of 21 stimulus pairs in the tactual modality while the visual subjects experienced the standard stimulus of each pair in the visual modality. Subjects in each subgroup were given two test sessions. One in which the comparative stimuli were presented in the same modality as the standard and in the other the comparative stimuli were presented in the alternate modality. These manipulations result in a two (social class) by two (modality of standard, stimulus) by two (intra- and inter-modality judgments) factorial design with repeated measures on the final factor. To the extent that cross modal transfer is a phenotypic indication of the Level II genotype LO-SES children would be predicted to be deficient in cross modal transfer relative to the HI-SES children. However, since the within modal judgments are indications of

Level I abilities which do not differ as a function of social class, no SES differences were predicted in the within modal ability.

Method

Forty children were selected from a population of children attending nursery school in the city of Lancaster, Pennsylvania. Four nursery schools cooperated in the study; two private church-related nursery schools and two public schools with preschool programs funded under Title I ESEA. All nursery age children in each school were given the Peabody Picture Vocabulary Test (PPVT) and had their social-economic status calculated by the Hollingshead Two Factor Index of Social Position (Myers & Bean, 1968). Ten children were assigned to each of four groups on the basis of SES scores with all groups matched for PPVT measures (see Table 4-1). These subject characteristics were analysed by a two (social class) by two (modality of standard) analysis of variance (see Appendices II, JJ, KK, LL, MM) with the only significant difference occurring as a main effect of social class in the Hollingshead social class index. This indicates that the matching procedures were successful and that two different social class groups were formed.

Stimulus objects were twelve 2 x 2 x ½ inch brown plastic cutouts fashioned after the nonsense shapes pictured by Millar (1971). These were divided into two sets of six objects with one set randomly assigned to the first testing session for each subject and the remaining set being presented in the second testing sessions for that subject.

These objects were present in pairs with the tactual presentations being given in a box 12" high x 18" wide x 24" deep. The box was open on one side so that E could place the object to be felt in S's hands both of which were inserted through a 4" x 12" cloth draped slit in the front of the box. The visually presented objects were placed upon the top of the box which was at, or slightly below, the eye level of Ss who were seated in a child-sized folding chair.

Three equal volume geometric solids (pyramid, sphere and cylinder) were employed as training stimuli to test each child's ability to make "same-different" judgments.

All Ss were administered the PPVT and the Hollingshead index approximately one month prior to testing. On the basis of those scores Ss were assigned to one of two groups within each social class level. All Ss were presented 21 pairs of stimuli (each member of the set of six paired with every other member of the set including itself) in which the first member of the pair was presented visually for one group and tactually for the second. This standard stimulus was presented for approximately 15 seconds and then removed for 15 seconds prior to presenting the second member of the pair. Ss were allowed to look at/feel this second

Table 4-1
Means and Standard Deviations of PPVT Results
and Social Class Index for all Groups

SES Level	Modality of Standard Presentation	CA	MA	IQ	PPVT Raw Score	SES Index
HI	Tactual	53.9 (5.76)	63.4 (8.67)	107.9 (5.35)	50.8 (4.48)	19.8 (6.41)
	Visual	53.8 (6.01)	64.0 (6.63)	107.9 (5.62)	51.3 (3.66)	20.1 (5.20)
LO	Tactual	52.2 (4.33)	58.7 (10.36)	105.1 (5.92)	48.1 (5.08)	59.7 (8.37)
	Visual	54.0 (4.04)	59.6 (8.92)	104.9 (6.10)	48.9 (4.84)	56.3 (11.38)

stimulus until they responded by deciding whether it was the same as, or different from, the first member of the pair.

All Ss in both the visual and tactual groups were given a second testing session one week after the first session. For one half of the subjects in each modality condition the first session involved intramodality judgments and the second consisted of intermodality comparisons while the order was reversed for the remainder of the S in each group.

Testing was accomplished by presenting the first member of the set and instructing the Ss to look at (or feel) it and to remember what it looks/feels like. That stimulus was removed in 15 seconds and after a 15 second interval, the second member of the pair was presented. After 15 seconds the S was asked whether it was the same as or different from the other one. Stimuli presented visually were held upright by E while S visually inspected them without being allowed to feel them. Tactual presentations were made by placing the upright object into the fingers of both hands of S which were inserted into the box. For intramodal testing both objects were presented to the S in the same modality while for intermodal testing the objects were presented via the appropriate method; visually on the top of the box and tactually inside the box.

The 21 pairs were presented in one of three predetermined random orders. Each stimulus set appeared equally frequently during the first and second testing sessions within and between all conditions. This eliminated any between-session transfer due to the specific stimulus objects yet controlled for any bias due to the specific sets. Upon completion of each testing session S was given his choice from a pool of 10 cent toys.

Prior to the initial presentations of the stimulus pairs a training session was given to provide instructions in the task and to guarantee that all Ss were capable of making correct "same-different" judgments. During this training period two of the equal volume stereometric forms (e.g., pyramid and cylinder) were handed to the S who was asked to make a "same-different" judgment. If the subject responded incorrectly he was corrected and given a brief explanation. If a correct response was made the S was asked to explain his response. This process was continued until S made 2 consecutive correct same and 2 correct different responses and could verbalize the form as the relevant dimension. S was then told he was going to play another game but this time he would see/feel only one thing at a time.

Results

The dependent measure was the number of correct responses in judging the 21 pairs of stimuli in each testing session (see Table 4-2).

Table 4-2

Means ($\overline{X}$), Standard Deviations (S.D.) and the Number of Subjects ($N_{.05}$) Performing at a Better than Chance Level in all Experimental Groups with t Values for Related Measures of Intramodal and Intermodal Errors

SES Level		Tactual Standard Intramodal		Tactual Standard Intermodal	Visual Standard Intramodal		Visual Standard Intermodal
HI SES	$\overline{X}$	5.6		7.5	1.5		4.8
	S.D.	1.26	t=3.45*	2.00	.97	t=6.35**	1.55
	$N_{.05}$	7		5	10		9
LO SES	$\overline{X}$	5.4		6.1	2.1		4.8
	S.D.	2.01	t=.74	1.97	.99	t=4.28*	1.63
	$N_{.05}$	7		8	10		8

* $df = 9$; $p<.01$
** $df = 9$; $p<.001$

An analysis of the error rate between the major experimental conditions was provided by a 2 (social class) by 2 (modality of standard) by 2 (intramodal vs. intermodal) analysis of variance with repeated measures on the intramodal versus intermodal comparisons (see Appendix NN). This analysis demonstrated no significant effect due to social class but significant main effects of the modality of the standard ($F = 57.79$; $df = 1,36$; $p < .005$) and the intra-intermodality condition ($F = 92.45$; $df = 1,36$, $p < .005$) with a significant interaction of modality of the standard with the inter-intramodality condition ($F = 14.45$, $df = 1,36$; $p < .025$). These results indicate that less errors are made when the standard is presented visually than tactually and that intramodal judgments are easier than the between modal judgments. Scheffé comparisons of the four cells formed by collapsing across the two social classes demonstrate that all comparisons are significant beyond the .05 level of confidence which indicates that the interaction occurs between groups all of which are different from each other. The interaction is probably due to the fact that the visual-intramodal condition is so much easier than any of the others.

To assess the absolute ability of the children to perform in this task the number of subjects who responded at a better than chance level in each group under both intramodal and intermodal presentations was calculated and is presented in Table 4-2. Given 21 choices under each condition and with a .5 probability of a correct response the expected mean correct responses would be 10.5 ($\hat{x} = np$) with an expected standard deviation of 2.29 ($S.D. = \sqrt{npq}$). A total of 15 or more correct responses produces a Z score which is significant at the 5% level. From the totals presented in Table 4-2 it is obvious that the only condition in which a large number of subjects were performing at a chance level is the HI-SES tactual-intermodal condition. Yet, Phi coefficients and simple Chi-square tests indicate no significant relationships between modality and SES level (phi = .07; chi square = .30) modality and intra-intermodal transfer (phi = .02; chi square = .03) or SES and intra-intermodal transfer (phi = .03; chi square = .07) in the number of Ss performing at an above chance level.

Finally, the difference in the Ss ability to make within and between modality judgments was analyzed with t tests for related measures (see Table 4-2). These tests indicate a significant decrement in intermodal responding relative to intramodal responding in all conditions except the tactual-visual condition for the LO-SES subjects.

Discussion

The analysis indicated that nursery school children were able to respond appropriately in making same-difference judgments of stimuli presented in either the visual or tactual sensory modality and in intermodal comparisons between these two modalities. Less errors were made in the intramodality than the intermodality conditions and in those judgments in which the standard stimuli were presented via the visual modality.

Most crucial for the present purposes was the lack of a deficiency in performance in the children from families of lower socioeconomic status. In fact, the only instance of any performance differences between the two SES levels appeared in the tactual-visual intermodality judgments and favored the LO-SES children. There were more LO-SES children performing at an above chance level than HI-SES children in this condition and the LO-SES children demonstrated no significant decrement in response rate on the tactual-visual intermodality judgment relative to their performance on the tactual-tactual intramodality judgments.

Unfortunately, since Jensen has provided no clearcut way of classifying tasks as measuring Level I or Level II abilities, the present data can be interpreted as indicating either a lack of SES differences in Level II abilities or the fact that CME tasks do not measure Level II abilities. In any instance, these data do not support the claim that children from lower socioeconomic status are deficient in cross modal transfer.

A distinction has been made (Blank & Bridger, 1964) between cross modal equivalence and cross modal concepts. The latter involve the transfer between modalities of a principle or concept not necessarily anchored to specific stimuli. It is possible that the cross modal equivalence tasks are not dependent upon Level II abilities. Support may be found for this limitation in studies which have demonstrated cross modal equivalence in monkeys (Rothblatt & Wilson, 1968) and retarded children (Smith & Tunick, 1969) but which have failed to produce any cross modal conceptual transfer of more abstract, mediational responses in these subject populations. However, given the procedural idiosyncracies of these studies and the facts that Tyrrell (1970) has demonstrated cross modal concepts in first grade children with mental ages identical to those of the retarded children of Smith and Tunick and that Pick, Pick and Thomas (1966) have demonstrated cross modal concept transfer in kindergarten age children, the distinction in processes underlying CME and CMC is not as clear as the procedural differences. In fact, Jensen (1969a) referring to the deficiencies in cross modal transfer related to nutritional deficiencies cites an article by Scrimsham (1968) which discusses deficiencies in cross modal equivalence, not cross modal concept, tasks.

It appears that corss modal equivalence tasks do require some central, symbolic, cognitive processing mechanism. Since the performance of LO-SES subjects in the present study was equivalent to the performance of the HI-SES subjects, it can be concluded that LO-SES subjects do not possess a lower level of ability to engage in those Level II processes.

General Discussion and Conclusions

The experiments discussed in this report were designed to assess the existence and extent of performance differences in nursery school children from different social classes. The particular tasks selected were similar in that adequate performance on each requires that the subject engage in some self-initiated, cognitive activity which intervenes between stimulus reception and response output. Tasks with this characteristic were selected since it is the authors' belief that the mediating process or symbolic activities demanded by these tasks represent important abilities for adequate and appropriate performance in the educational world which will soon be presented to the child. In fact, another more meaningful word which refers to these abilities would be "thinking". Unfortunately, "thinking" may have too much meaning and a specific description of the particular "thinking" being investigated is required.

The particular cognitive abilities investigated included: The ability to focus attention on selected, relevant aspects of the stimulus world and to generalize this to new situations; the ability to organize information which is to be memorized so that the amount which will be remembered is maximized; and, the ability to code information presented through one sensory modality so that the information can be utilized in a different modality. If a child were able to pay attention to the appropriate parts of his world, could encode the information presented through vision, touch, audition, etc., and could somehow commit that information to memory for total recall, he would certainly succeed in school and would be an unusual adult. In fact, the educational system which the Western world has developed seems to make the implicit assumption that all children do have these abilities and will use them in the classroom. However, educators have long recognized the fact that all children do not possess these abilities to an equal extent, or at least don't behave in the classroom as if they did. The most popular and perhaps the most successful attempt to measure indirectly the extent to which any child does possess these, and other similar, abilities is the intelligence test. These tests have, at various times, been assumed to measure the ability of an individual to think rationally, or even the potential of an individual to do so. Now, however, it is generally recognized that the intelligence test, and its derived score, the IQ, is really best considered as an index of present performance on a particular, relatively restricted, range of material. Performance on standardized tests of general intelligence do relate relatively well to performance in academic subjects in which it is reasonable to suppose one should have to "think" in order to do well. However, the relationship between intelligence test results and academic performance is not perfect and it is obvious that some factor, or factors, not measured by intelligence tests are important for academic performance.

Most recently, Jensen (1969b) has proposed that levels of performance on a wide range of tasks cannot be thought of as being determined by one type of underlying ability which each individual possesses to one degree or another. Jensen cites, for example, the relatively common observation that some children who are not "doing well" on the common indices of "intellectual ability" (i.e., grades and test results) perform exceptionally well at "non-academic" tasks (e.g., learning the names of 25-40 new classmates by the end of the first day of school). These and similar observations caused Jensen and several of his colleagues to initiate a program of research to investigate these anomalies of performance. The method of investigation selected by Jensen was to measure ability on a set of tasks directly rather than indirectly. Tasks such as digit recall, serial learning, paired-associate learning and free recall were presented to children varying in age, intelligence test performance and social class. In summarizing the results of these investigations Jensen (1969b) concluded that performance on these tasks was highly related to IQ in middle class children but that there appeared little or no relation between the two measures of ability in lower class children. In other words, the IQ scores of the lower social class children did not relate to performance on the learning tasks; the children with low IQ did as well on the learning tasks as the high IQ children.

To account for these, and other results, Jensen hypothesized that each task (including the intelligence tests) must be analyzed in terms of two different, independent factors; its "cultural loading" (from culture free to culture loaded) and the complexity of learning (from simple associative learning to abstract problem solving and conceptual learning).

The culture loading of any task is its heritability (h^2) or the extent to which performance on the task is genetically determined (if a task is culture free ($h^2=1$), performance on that task is not affected by cultural or environmental factors). Jensen (1969a) estimates the heritability of intelligence as measured by the Stanford Binet as approximately .75 with the heritability of scholastic attainment estimated to be considerably lower.

The second dimension, complexity of learning tasks, is the more important for our purposes. This dimension represents a continuum which ranges from tasks which are predominantly measures of rote memory (such as digit span) through tasks which demand more self initiated activity of the child. Tasks in which ".. the subject must spontaneously bring more ... covert 'mental' activity (discrimination, generalization, verbal mediation, deduction, induction, and hypothesis testing) to bear on the task in order to perform successfully (Jensen, 1969b, p. 30)." Jensen hypothesizes that tasks ranging from associative to conceptual learning are the phenotypic expression of two functionally dependent but genetically (structurally) independent mental processes.

labelled simply Level I and Level II. Most noteworthy is the postulation that Level I and Level II abilities are distributed differently in upper and lower social classes while Level I abilities are distributed equally in all social classes, Level II abilities are distributed around a higher mean in higher social classes. According to Jensen (1969b) three hypotheses account for their results. These are: "...(a) the genotypic independence of Level I and Level II processes, (b) the functional dependence of Level II upon Level I, and (c) the differential distribution of individual differences in Level I and Level II genotypes in upper and lower social classes...(p. 31)."

This analysis points to a serious shortcoming of the general tests of abilities currently available in education. Tests have not yet been developed which distinguish between the functional levels of Level I and Level II abilities. In fact, two quite different individuals (in terms of Level I and Level II abilities) may appear quite similar in measures of overall intelligence but be quite different in their performance on a range of scholastic tasks which acquire different amounts of the two abilities.

The tasks of the present report are all on the Level II end of the "Level of complexity of learning continuum", and are, therefore, excellent tasks for testing Jensen's differential distribution hypothesis. They provide a particularly crucial test since the children from varying levels of social class were selected such that there were no differences in performance on a test of general intellectual ability, the Peabody Picture Vocabulary Test (PPVT). The PPVT, like most intelligence tests, does not distinguish between Level I and Level II abilities and it is quite possible that any particular range of performance on this test can be achieved by children with different amounts of the two abilities. This would be particularly true if the performance range were relatively restricted and within the average ranges. Therefore, children from the lower social classes who perform at the average for their age norms may possess above average Level I abilities which would compensate for their decreased Level II abilities relative to the children from higher social classes. If this were true, the differences in Level II abilities of these children would appear as performance differences in the experimental tasks of the present investigations. Obviously, given this logic, a lack of difference in performance on the experimental tasks would not allow the conclusion that Level II abilities are not differentially distributed in the two social classes since controlling for performance on the PPVT may have equated Level II abilities in the sample of children selected. However, given the raw scores of the children on the PPVT, the level of abstraction of the vocabulary items identified by the children is extremely low and are probably tapping Level I abilities to a large extent.

The results of the three experiments do not conform to the simple notion of social classes differences in Level II abilities.

Yet, there were significant differences in behavior between the children of high and low social class. These differences were the type which would normally be explained by postulating differences in intellectual ability; an explanation which is inappropriate since intellectual performance was not allowed to vary.

In the first experiment, the LO-SES children did not acquire the first extradimensional shift as rapidly as the HI-SES children, and this performance decrement continued for successive problems. The specific process postulated to account for this deficiency was decreased ability to learn to "pay attention" to the appropriate physical dimension. This deficiency has been demonstrated to be a characteristic of mentally retarded children and was assumed to be determined by intelligence. In the present experiment the subject characteristic associated with the deficiency was not intelligence but rather social class; a result which is consistent with the notions of Jensen.

However, following a series of such transfer problems the performance of the LO-SES and HI-SES children converged at a high level. This converging indicates a large improvement in the ability of all children to acquire the appropriate attentional response. To the extent that this represents an increase in a basic ability which is a Level II process, the Level II ability is subject to environmental effects. This would represent a placement on the "cultural loading" continuum on the culture loaded end of the continuum with a low heritability index. If that were true, then the observed differences in distribution of the ability in different social classes could be altered by "engineering" the environment in an appropriate manner. Future investigations should certainly be conducted to assess the heritability of this ability and to specify those environmental events important for maximizing the ability.

In the second experiment the primary cognitive process being investigated was that of organizing stimulus input for maximal recall. This ability was not shown to differ greatly in children from different social classes. However, the LO-SES children did produce a recall list which possessed a large number of inappropriate items. This characteristic is not necessarily a deficiency in some structurally determined process, but is more appropriately considered a strategy for recall which is quite efficient in producing a high recall level as long as it is not important (or damaging) to have a concomitantly high number of errors. Interestingly enough, this unedited recall output may result in a more accurate estimate of the limits of a subject's recall ability. This, too, is subject to further experimental investigation and may provide a phenotype measure which more closely reflects the genotypical process level.

Finally, the third experiment demonstrated no meaningful social class differences in cross modal transfer. The task, although demanding some central coding mechanism which transforms the modality specific stimulus into one with multi-modal recognizability, may not be appropriate for assessing the presence of potential ability differences. A more appropriate task might be one which demands the application of a more abstract concept not anchored to any specific stimulus objects. Future research should employ tasks which would demand the cross modal concepts of Blank and Bridger (1966) rather than cross modal equivalence.

One final comment involves the relevance of the current results to Jensen's recommendation that educational researchers discover and devise instructional techniques which minimize the need for high amounts of Level II for academic success. The present authors agree that educators must capitalize on whatever abilities the child possesses and, therefore, agree that children with high degrees of Level I ability should have the opportunity to succeed educationally by being able to utilize that ability to acquire the basic academic skills. However, to rely exclusively upon Level I abilities limits the effective cognitive operations which a child may acquire and, therefore, limits the ultimate academic and intellectual achievements. The present data question the implicit assumption that Level II abilities are high in heritability and suggests the opposite may be the case. Therefore, it is imperative that educational researchers assess the culture loading of various tasks which require Level II type abilities and specify the environmental determinants of those processes. Implicit in that suggestion is the notion of several, largely independent, cognitive processes each of which may have idiosyncratic environmental determinants and a different heritability index.

References

Bereiter, C. & Engelmann, S. Teaching disadvantaged children in the preschool. Engelwood Cliffs: Prentice-Hall, 1966.

Birch, H. G. & Lefford, A. Intersensory development in children, Monographs of Society for Research in Child Development, 1963, 28(Whole No. 89).

Blank, M. & Bridger, W. H. Conceptual cross-modal transfer in deaf and hearing children, Child Development, 1966, 37, 29-38.

Bousfield, A. K. & Bousfield, W. A. Measurement of clustering and of sequential constancies in repeated free recall, Psychological Reports, 1966, 19, 935-942.

Bousfield, W. A., Puff, C. R. & Cowan, T. M. The development of constancies in sequential organization during repeated free recall, Journal of Verbal Learning and Verbal Behavior, 1964, 3, 489-495.

Bousfield, W. A. & Rosner, S. R. Free vs uninhibited recall, Psychonomic Science, 1970, 20, 75-76.

Connors, C. K., Schuette, C. & Goldman, Ann Informational analysis of intersensory communication in children of different social class, Child Development, 1967, 38, 251-267.

Dalrymple-Alford, E. C. The measurement of clustering in free recall, Psychological Bulletin, 1970, 73, 32-34.

Deutsch, M. P. The disadvantaged child and the learning process. In Passow, A. H. (Ed.), Education in depressed areas. New York: Bureau of Publications, Teachers College, Columbia University, 1963.

Fowler, W. Cognitive learning in infancy and early childhood, Psychological Bulletin, 1962, 59, 116-152.

Glasman, L. D. A social-class comparison of conceptual processes in children's free recall. Unpublished doctoral dissertation, University of California, Berkeley, 1968.

Gorden, E. W. & Wilkerson, D. A. Compensatory education for the disadvantaged. New York: College Entrance Examination Board, 1966.

Hollingshead, A. B. Social class and mental illness: A community study. New York: John Wiley & Sons, 1958.

Hunt, J. McV., The psychological basis for using preschool enrichment as an antidote for cultural deprivation, Merrill-Palmer Quarterly, 1964, 10, 209-248.

Inhelder, B. & Piaget, J. The growth of logical thinking. New York: Basic Books, 1958.

Jensen, A. R. Learning abilities in Mexican-American and Anglo-American children, California Journal of Educational Research, 1961, 12, 147-159.

Jensen, A. R. Learning in the preschool years. In Hartrup, W. W. & Smothergill, N. L. (Eds.), The young child. Washington: National Association for the Education of Young Children, 1967.

Jensen, A. R. Social class and verbal learning. In Deutsch, M., Katz, I., & Jensen, A. R. (Eds.), Social class, race, and psychological development. New York: Holt, Rinehart, and Winston, Inc., 1968.

Jensen, A. R. Patterns of mental ability and socioeconomic status. Proceedings of the National Academy of Sciences, 1968, 60, 1330-1337.

Jensen, A. R. How much can we boost IQ and scholastic achievement?, Harvard Educational Review, 1969a, 39, 1-123.

Jensen, A. R. Intelligence, learning, ability and socioeconomic status, Journal of Special Education, 1969b, 31(1), 23-35.

Jensen, A. R. The role of verbal mediation in mental development, Journal of Genetic Psychology, 1971, 118, 39-70.

Jensen, A. R. & Frederiksen, J. Social-class differences in free recall learning, Journal of Educational Psychology, in press.

Kendler, H. H. & Kendler, T. S. Vertical and horizontal processes in problem solving, Psycholgical Review, 1962, 69, 1.

Klaus, R. A. & Gray, S. W. The early training project for disadvantaged children: A report after five years, Monograph for the Society for Research in Child Development, 1968, 33 (4, Serial No. 120), 1-66.

Leeper, S. H., Kales, R. J., Skipper, D. S. & Witherspoon, R. L. Good schools for young children. New York: Macmillan, 1968.

Mandler, G. Organization and memory. In Spence, K. W. & Spence, J. T. (Eds.), The psychology of learning and motivation. New York: Academic Press, 1967, 1, 327-372.

Mandler, G. & Dean, P. J. Seriation: Development of serial order in free recall, Journal of Experimental Psychology, 1969, 81, 207-215.

Millar, S. Visual and haptic cue utilization by preschool children: The recognition of visual and haptic stimuli presented separately and together, Journal of Experimental Child Psychology, 1971, 12, 88-94.

Montessori, M. The Montessori method. New York: Schocken Books, 1964. (Translated from the Italian, 1912)

Myers, J. K. & Bean, L. L. A decade later: A follow up of social class and mental illness. New York: John Wiley & Sons, 1968.

Passow, A. H. Education in depressed areas. New York: Bureau of Publications, Teachers College, Columbia University, 1963.

Pick, A. D., Pick, H. L. & Thomas, M. L. Cross-modal transfer and improvement in form discrimination, Journal of Experimental Psychology, 1966, 71, 279-288.

Puff, C. R. An investigation of two forms of organization in free recall, Journal of Verbal Learning and Verbal Behavior, 1970, 9, 720-724.

Rapier, J. L. The learning abilities of normal and retarded children as a function of social class, Journal of Educational Psychology, 1968, 59, 102-119.

Roenker, D. L., Thompson, C. P. & Brown, S. C. Comparison of measures for the estimation of clustering in free recall, Psychological Bulletin, 1971, 76, 45-48.

Rothblatt, L. A. & Wilson, W. A. Intradimensional and extra-dimensional in the monkey within and across sensory modalities, Journal of Comparative and Physiological Psychology, 1968, 66, 539-543.

Scrimshaw, N. S. Infant malnutrition and adult learning. Saturday Review, March 16, 1968.

Shepp, B. E. & Turrisi, F. D. Learning and transfer of mediating respnses in discrimination learning. In Ellis, N. R. (Ed.), Research in mental retardation. New York: Academic Press, 1966.

Smith, J. & Tunick, J. Crossmodal transfer of a discrimination by retarded children, Journal of Experimental Child Psychology, 1969, 7, 274-281.

Spitz, H. H. The role of input organization in the learning and memory of mental retardates. In Ellis, N. R. (Ed.), Research in mental retardation. New York: Academic Press, 1966.

Tulving, E. Subjective organization in free recall of "unrelated" words, Psychological Review, 1962, 69, 344-354.

Tulving, E. Theoretical issues in free recall. In Dixon, T. R. & Horton, D. L. (Eds.), Verbal behavior and general behavior theory. New Jersey: Prentice-Hall, Inc., 1968.

Tyrrell, D. J. Intradimensional and extradimensional shifts and cross-modal transfer in children. Paper presented at the meetings of the Eastern Psychological Association, Atlantic City, April 1970.

Underwood, B. J. & Schulz, R. W. Meaningfulness and verbal learning. Chicago, Illinois: Lippincott, 1960.

Weikart, D. P. Perry preschool project progress report. Ypsilanti (Michigan) Public Schools, 1964.

Whiteman, M. & Deutsch, M. Social disadvantage as related to intellective and language development. In Deutsch, M., Katz, I., & Jensen, A. R. (Eds.), Social class, race and psychological development. New York: Holt, Rinehart & Winston, 1968.

Zeaman, D. & House, B. J. An attention theory of retardate discrimination learning. In Ellis, N. R. (Ed.), Handbook of Mental Deficiency. New York: McGraw-Hill, 1963.

Appendix A

Summary of Analysis of Variance for MA Score on the PPVT

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs Lo)	312.5	1	1.37	<.20
B (Shift: ID vs ED)	162.0	1	.71	<.20
A x B	.5	1	.002	<.20
Error	228.2	28		

Appendix B

Summary of Analysis of Variance for CA

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs Lo)	.12	1	.003	<.20
B (Shift: ID vs ED)	.12	1	.003	<.20
A x B	2.25	1	.055	<.20
Error	40.5	28		

Appendix C

Summary of Analysis of Variance for IQ on the PPVT

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs Lo)	270.3	1	.867	<.20
B (Shift: ID vs ED)	166.6	1	.534	<.20
A x B	463.1	1	1.485	<.20
Error	311.8	28		

Appendix D

Summary of Analysis of
Variance for PPVT Raw Score

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs Lo)	98.00	1	.966	<.20
B (Shift: ID vs ED)	98.00	1	.966	<.20
A x B	1.125	1	.011	<.20
Error	101.42	28		

Appendix E

Summary of Analysis
of Variance for SES Index

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs Lo)	15400.13	1	176.56	<.001
B (Shift: ID vs ED)	50.00	1	.57	<.20
A x B	.5	1	.000	<.20
Error	87.223	28		

Appendix F

Summary of Analysis for Errors
to Criterion on Original Learning

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs Lo)	180.5	1	.128	>.20
B (Shift: ID vs ED)	986.00	1	.700	>.20
A x B	220.50	1	.156	>.20
Error	1407.59	28		

Appendix G

Summary of Analysis of Variance for Errors
To Criterion on the First Transfer Problem

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs Lo)	1225.13	1	2.726	> .10
B (Shift: ID vs ED)	4851.12	1	10.796	< .005
A x B	2178.00	1	4.847	< .05
Error	449.33	28		

Appendix H

Summary of Analysis of Variance
for Errors on First Six Transfer Problems

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs Lo)	5002.083	1	4.69	< .05
B (Shift: ID vs ED)	7777.521	1	7.30	< .025
AB	776.021	1	.73	
Error (between Ss)	1065.454	28		
C (problems)	618.496	5	1 82	
AC	238.196	5	.70	
BC	833.933	5	2.46	< .05
ABC	367.983	5	1.09	
Error (within Ss)	339.045	140		

Appendix I

Summary of Analysis of Variance for Logarithm Transformation of Errors Over the First Six Transfer Problems

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs Lo)	3.960	1	7.05	< .025
B (Shift: ID vs ED)	4.810	1	8.57	< .01
AB	.510	1	.91	
Errors (between Ss)	.562	28		
C (problems)	.685	5	3.28	< .01
AC	.054	5	.26	
BC	.635	5	3.04	< .01
ABC	.200	5	.96	
Errors (within Ss)	.209	140		

Appendix J

Summary of Analysis for
Errors on All Transfer Problems

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs Lo)	3400.10	1	4.72	< .05
B (Shift: ID vs ED)	6966.92	1	9.67	< .005
AB	840.72	1	1.17	
Error (between Ss)	720.50	28		
C (problems)	1084.91	10	4.40	< .005
AC	294.67	10	1.20	
BC	637.81	10	2.59	< .005
ABC	209.33	280	.85	
Error (within Ss)	246.30			

Appendix K

Summary of Analysis of Variance for Logarithm Transformation of Error Score Over All Transfer Problems

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs Lo)	2.78	1	5.19	<.05
B (Shift: ID vs ED)	6.52	1	12.16	<.001
AB	1.35	1	2.52	
Error (between Ss)	.54			
C (problems)	.88	10	5.38	<.001
AC	.16	10	1.01	
BC	.38	10	2.31	<.025
ABC	.14	10	.88	
Error (within Ss)	.16	230		

Appendix 3.

Summary of Analysis of Variance for
Errors to Criterion on the Last Transfer Problem

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs Lo)	.28	1	.14	
B (Shift: ID vs ED)	16.53	1	8.07	<.01
A x B	7.03	1	3.43	
Error	2.05	28		

Appendix M

Stimulus Lists Used in this Study

List No. 1	List No. 2	List No. 1	List No. 2
French Fries	Hammer	Bridge	Hammer
Hamburger	Shovel	Ice Cream Cone	Iron
Ice Cream Cone	Wrench	Watch	Tree
Bus	Apple	Car	Barn
Car	Grapes	Flag	Grapes
Truck	Oranges	Telephone	Letter
Elephant	Bed	Drum	Airplane
Lion	Chair	Lamp	Chair
Monkey	Table	Monkey	Whistle
Dress	Fireman	Glasses	Dog
Pants	Mailman	Pants	Fireman
Shoes	Policeman	Rake	Shoes

Appendix N

Summary of Analysis of Variance for
Mean Number of Items Recalled Correctly

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs. Lo)	.04	1	.00	
B (Type of List: C vs. NC)	178.90	1	17.25	<.001
C (List No.: 1 vs. 2)	4.90	1	.47	
A x B	29.04	1	2.80	.10>p>.05
A x C	.93	1	.09	
B x C	20.45	1	1.97	
A x B x C	100.15	1	9.66	<.005
Error (Between Ss)	10.37	64		
D (Trials)	86.96	5	50.43	<.001
A x D	3.25	5	1.89	
B X D	4.27	5	2.48	<.05
C x D	.95	5	.55	
A x B x D	.30	5	.17	
A x C x D	.98	5	.57	
B x C x D	.21	5	.12	
A x B x C x D	.60	5	.35	
Error (Within Ss)	1.72	320		

Appendix O

Summary of Analysis of Variance for
Total Responses Produced

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs. Lo)	80.95	1	2.36	
B (Type of List: C vs. NC)	204.19	1	5.94	<.025
C (List No.: 1 vs. 2)	93.52	1	2.72	
A x B	111.02	1	3.23	.10>p>.05
A x C	6.02	1	.18	
B x C	95.39	1	2.78	
A x B x C	105.02	1	3.06	.10>p>.05
Error (Between Ss)	34.36	64		
D (Trials)	275.00	5	43.51	<.001
A x D	16.70	5	2.64	=.025
B x D	5.05	5	.80	
C x D	8.71	5	1.38	
A x B x D	.93	5	.15	
A x C x D	4.68	5	.74	
B x C x D	6.74	5	1.07	
A x B x C x D	10.57	5	1.67	
Error (Within Ss)	6.32	320		

Appendix P

Summary of Analysis of Variance for
Total Errors Transformed by Log (x+1)

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs. Lo)	1.44	1	7.66	<.01
B (Type of List: C vs. NC)	.01	1	.07	
C (List No.: 1 vs. 2)	.75	1	3.98	<.05
A x B	.26	1	1.33	
A x C	.07	1	.40	
B x C	.47	1	2.52	
A x B x C	.13	1	.67	
Error (Between Ss)	.19	64		
C (Trials)	1.17	5	22.77	<.001
A x D	.07	5	1.42	
B x D	.03	5	.54	
C x D	.04	5	.71	
A x B x D	.01	5	.25	
A x C x D	.03	5	.58	
B x C x D	.12	5	2.26	<.05
A x B x C x D	.16	5	3.12	<.01
Error (Within Ss)	.05	320		

Appendix Q

Summary of Analysis of Variance for
Total Errors

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs. Lo)	75.84	1	5.01	$<.05$
B (Type of List: C vs. NC)	1.95	1	.13	
C (List No.: 1 vs. 2)	61.50	1	4.06	$<.05$
A x B	22.69	1	1.50	
A x C	8.61	1	.57	
B x C	22.69	1	1.50	
A x B x C	.39	1	.03	
Error (Between Ss)	15.13	64		
D (Trials)	62.50	5	14.67	$<.001$
A x D	8.78	5	2.06	$.10>p>.05$
B x D	.68	5	.16	
C x D	5.88	5	1.38	
A x B x D	1.17	5	.27	
A x C x D	1.67	5	.39	
B x C x D	8.29	5	1.94	$.10>p>.05$
A x B x C x D	7.27	5	1.71	
Error (Within Ss)	4.26	320		

Appendix R

Summary of Analysis of Variance for
Correct Recall as a Percentage of Total Words Produced

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs. Lo)	5213.13	1	9.65	<.005
B (Type of List: C vs. NC)	260.66	1	.48	
C (List No.: 1 vs. 2)	2424.46	1	4.49	<.05
A x B	559.67	1	1.04	
A x C	554.63	1	1.03	
B x C	848.43	1	1.57	
A x B x C	17.20	1	.03	
Error (Between Ss)	540.05	64		
D (Trials)	1746.68	1	10.75	<.001
A x D	123.10	5	.76	
B x D	138.36	5	.85	
C x D	73.36	5	.45	
A x B x D	37.73	5	.23	
A x C x D	97.79	5	.60	
B x C x D	460.96	5	2.84	<.025
A x B x C x D	372.99	5	2.29	<.05
Error (Within Ss)	162.54	320		

Appendix S

Brief Definitional Formulas for the Clustering Measures Used in this Study

Common Terminology

O(SCR) = Number of observed stimulus category repetitions.
E(SCR) = Number of repetitions expected by chance.
Max(SCR) = Maximum possible SCR value for a given protocol.
Min(SCR) = Minimum possible SCR value for a given protocol.
N = Total number of words recalled.

Measure	Investigator(s)	Formula
O-E(SCR) Difference	Bousfield & Bousfield (1966)	O-E(SCR) = O(SCR) - E(SCR)
Ratio of Clustering	Puff (1970)	$\text{Ratio} = \frac{O(SCR)}{Max(SCR)}$
Adjusted Ratio of Clustering	Roenker, Thompson, & Brown (1971)	$ARC = \frac{O(SCR) - E(SCR)}{Max(SCR) - E(SCR)}$
D_A Index	Dalrymple-Alford (1970)	$D_A = \frac{O(SCR) - E(SCR)}{Max(SCR) - Min(SCR)}$

Appendix T

Summary of Analysis of Variance for
Mean Observed-Expected SCR Difference Scores
for Categorized Lists

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs. Lo)	10.69	1	2.24	
B (List No.: 1 vs. 2)	2.46	1	.52	
A x B	13.18	1	2.77	
Error (Between Ss)	4.76	32		
C (Trials)	17.28	5	11.88	<.001
A x C	.44	5	.30	
B x C	1.17	5	.81	
A x B x C	.39	5	.27	
Error (Within Ss)	1.45	160		

Appendix U

Summary of Analysis of Variance for
Observed Clustering as a Proportion of Maximum Possible
(Ratio SCR Scores)

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs. Lo)	.07	1	.56	
B (List No.: 1 vs. 2)	.00	1	.01	
A x B	.21	1	1.57	
Error (Between Ss)	.13	29		
C (Trials)	.17	5	3.45	<.005
A x C	.05	5	1.11	
B x C	.03	5	.66	
A x B x C	.06	5	1.12	
Error (Within Ss)	.05	45		

Appendix V

Summary of Analysis of Variance for Mean Adjusted Ratio of Clustering Scores (ARC) for the Categorized Lists

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs. Lo)	.09	1	.16	
B (List No.: 1 vs. 2)	.17	1	.31	
A x B	1.39	1	2.52	
Error (Between Ss)	.55	29		
C (Trials)	.85	5	4.54	<.001
A x C	.26	5	1.39	
B x C	.13	5	.72	
A x B x C	.11	5	.60	
Error (Within Ss)	.19	145		

Appendix W

Summary of Analysis of Variance for
D_A Index Clustering Scores for the Categorized Lists

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs. Lo)	.01	1	.04	
B (List No.: 1 vs. 2)	.07	1	.35	
A x B	.53	1	2.80	
Error (Between Ss)	.19	29		
C (Trials)	.35	5	5.94	<.005
A x C	.06	5	1.07	
B x C	.06	5	.99	
A x B x C	.03	5	.52	
Error (Within Ss)	.06	145		

Appendix X

Brief Definitional Formulas for the Subjective Organization Measures Used in this Study

Common Terminology

O(ITR) = Number of observed intertrial repetitions.
E(ITR) = Number of repetitions expected by chance.
Max(ITR) = Maximum possible ITR value for a pair of protocols.

Measure	Investigator(s)	Formula
O-E(ITR) Difference	Bousfield, Puff, and Cowan (1964)	$O\text{-}E(ITR) = O(ITR) - E(ITR)$
Ratio of ITR	Puff (1970)	$\text{Ratio} = \frac{O(ITR)}{Max(ITR)}$
Adjusted Ratio of Subjective Organization	New in the present study	$ARSO = \frac{O(ITR)-E(ITR)}{Max(ITR)-E(ITR)}$

Notes:

All measures were duplicated for both unidirectionaly (U) and bidirectionally (B) defined units of ITR.

The O-E(ITR) and ARSO measures were, in addition, found using E(ITR) values found with the standard formula and with values, $E^1(ITR)$, estimated from statistical permutations.

Appendix Y

Summary of Analysis of Variance for O-E(ITR$_U$) Difference Scores

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs. Lo)	.00	1	.00	
B (Type of List: C vs. NC)	12.55	1	6.56	<.025
C (List No.: 1 vs. 2)	.25	1	.13	
A x B	1.62	1	.85	
A x C	.07	1	.04	
B x C	.28	1	.15	
A x B x C	.06	1	.03	
Error (Between Ss)	1.91	64		
D (Trial Pairs)	11.38	4	9.50	<.001
A x D	.30	4	.25	
B x D	3.38	4	2.82	<.05
C x D	1.66	4	1.38	
A x B x D	1.08	4	.90	
A x C x D	.38	4	.31	
B x C x D	1.73	4	1.44	
A x B x C x D	1.39	4	1.16	
Error (Within Ss)	1.20	256		

Appendix Z

Summary of Analysis of Variance for
O-E'(ITR$_U$) Difference Scores

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs. Lo)	.20	1	.15	
B (Type of List: C vs. NC)	16.26	1	12.33	<.001
C (List No.: 1 vs. 2)	.77	1	.59	
A x B	.07	1	.05	
A x C	.05	1	.04	
B x C	1.08	1	.82	
A x B x C	.26	1	.20	
Error (Between Ss)	1.32	64		
			5.53	
D (Trial Pairs)	5.90	4	.13	<.001
A x D	.14	4	4.03	
B x D	4.29	4	.78	<.005
C x D	.83	4	.82	
A x B x D	.87	4	.24	
A x C x D	.25	4	1.57	
B x C x D	1.68	4	.89	
A x B x C x D	.95	4		
Error (Within Ss)	1.06	256		

Appendix AA

Summary of Analysis of Variance for
O-E(ITR_B) Difference Scores

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs. Lo)	2.02	1	.66	
B (Type of List: C vs. NC)	19.32	1	6.35	<.025
C (List No.: 1 vs. 2)	4.41	1	1.45	
A x B	3.06	1	1.01	
A x C	.07	1	.02	
B x C	2.04	1	.67	
A x B x C	.41	1	.14	
Error (Between Ss)	3.04	64		
D (Trial Pairs)	17.84	4	14.35	<.001
A x D	.36	4	.29	
B x D	4.39	4	3.53	<.01
C x D	1.21	4	.98	
A x B x D	2.15	4	1.73	
A x C x D	.43	4	.35	
B x C x D	1.60	4	1.29	
A x B x C x D	2.14	4	1.72	
Error (Within Ss)	1.24	256		

Appendix BB

Summary of Analysis of Variance for
O-E'(ITR_B) Difference Scores

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs. Lo)	.03	1	.02	
B (Type of List: C vs. NC)	27.56	1	17.35	<.001
C (List No.: 1 vs. 2)	.04	1	.03	
A x B	.04	1	.03	
A x C	.00	1	.00	
B x C	1.42	1	.89	
A x B x C	.48	1	.30	
Error (Between Ss)	1.59	64		
D (Trial Pairs)	6.07	4	5.86	<.001
A x D	.73	4	.71	
B x D	2.69	4	2.60	<.05
C x D	.84	4	.81	
A x B x D	1.79	4	1.73	
A x C x D	.40	4	.38	
B x C x D	.37	4	.36	
A x B x C x D	1.36	4	1.32	
Error (	1.03	256		

Appendix CC

Summary of Analysis of Variance for Observed ITR_U Scores as a Proportion of Maximum Possible Values

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs. Lo)	.00	1	.01	
B (Type of List: C vs. NC)	.01	1	.15	
C (List No.: 1 vs. 2)	.49	1	.84	
A x B	.02	1	.31	
A x C	.12	1	2.06	
B x C	.01	1	.22	
A x B x C	.07	1	1.26	
Error (Between Ss)	.06	64		
D (Trial Pairs)	.20	4	3.88	<.01
A x D	.01	4	.29	
B x D	.04	4	.71	
C x D	.04	4	.77	
A x B x D	.05	4	.96	
A x C x D	.02	4	.38	
B x C x D	.01	4	.23	
A x B x C x D	.13	4	2.51	<.05
Error (Within Ss)	.05	256		

Appendix DD

Summary of Analysis of Variance for
Observed ITR_B Scores as a Proportion of
Maximum Possible Values

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs. Lo)	.02	1	.21	
B (Type of List: C vs. NC)	.02	1	.31	
C (List No.: 1 vs. 2)	.01	1	.06	
A x B	.04	1	.52	
A x C	.14	1	1.83	
B x C	.02	1	.25	
A x B x C	.27	1	3.51	$.10>p>.05$
Error (Between Ss)	.08	64		
D (Trial Pairs)	.23	4	4.01	$<.005$
A x D	.03	4	.58	
B x D	.17	4	3.03	$<.025$
C x D	.10	4	1.76	
A x B x D	.12	4	2.04	$.10>p>.05$
A x C x D	.09	4	1.52	
B x C x D	.01	4	.25	
A x B x C x D	.16	4	2.77	$<.05$
Error (Within Ss)	.06	256		

Appendix EE

Summary of Analysis of Variance for
Adjusted Ratio of Subjective Organization ($ARSO_U$) Scores

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs. Lo)	.01	1	.15	
B (Type of List: C vs. NC)	.00	1	.03	
C (List No.: 1 vs. 2)	.07	1	1.06	
A x B	.01	1	.13	
A x C	.10	1	1.55	
B x C	.03	1	.48	
A x B x C	.04	1	.61	
Error (Between Ss)	.06	64		
D (Trial Pairs)	.31	4	4.80	<.001
A x D	.02	4	.35	
B x D	.04	4	.62	
C x D	.05	4	.80	
A x B x D	.06	4	1.00	
A x C x D	.02	4	.28	
B x C x D	.01	4	.21	
A x B x C x D	.16	4	2.54	<.05
Error (Within Ss)	.06	256		

Appendix FF

Summary of Analysis of Variance for
Adjusted Ratio of Subjective Organization ($ARSO_U'$) Scores

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs. Lo)	.00	1	.04	
B (Type of List: C vs. NC)	.02	1	.22	
C (List No.: 1 vs. 2)	.12	1	1.74	
A x B	.02	1	.34	
A x C	.14	1	2.00	
B x C	.13	1	1.81	
A x B x C	.02	1	.26	
Error (Between Ss)	.07	64		
D (Trial Pairs)	.26	4	3.71	<.01
A x D	.03	4	.40	
B x D	.05	4	.68	
C x D	.07	4	.96	
A x B x D	.10	4	1.40	
A x C x D	.01	4	.21	
B x C x D	.02	4	.22	
A x B x C x D	.16	4	2.35	.10>p>.05
Error (Within Ss)	.07	256		

Appendix GG

Summary of Analysis of Variance for
Adjusted Ratio of Subjective Organization ($ARSO_B$) Scores

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs. Lo)	.05	1	.38	
B (Type of List: C vs. NC)	.01	1	.05	
C (List No.: 1 vs. 2)	.08	1	.68	
A x B	.09	1	.72	
A x C	.12	1	1.01	
B x C	.39	1	3.16	$.10>p>.05$
A x B x C	.45	1	3.72	$.10>p>.05$
Error (Between Ss)	.12	64		
D (Trial Pairs)	.48	4	4.07	<.005
A x D	.11	4	.95	
B x D	.40	4	3.43	<.01
C x D	.10	4	.83	
A x B x D	.14	4	1.21	
A x C x D	.20	4	1.70	
B x C x D	.15	4	1.28	
A x B x C x D	.21	4	1.78	
Error (Within Ss)	.12	256		

Appendix HH

Summary of Analysis of Variance for
Adjusted Ratio of Subjective Organziation ($ARSO_B$') Scores

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs. Lo)	.02	1	.17	
B (Type of List: C vs. NC)	.14	1	1.14	
C (List No.: 1 vs. 2)	.06	1	.48	
A x B	.03	1	.21	
A x C	.13	1	1.07	
B x C	.27	1	2.17	
A x B x C	.21	1	1.68	
Error (Between Ss)	.12	64		
D (Trial Pairs)	.46	4	3.70	<.01
A x D	.05	4	.41	
B x D	.15	4	1.21	
C x D	.10	4	.76	
A x B x D	.16	4	1.29	
A x C x D	.12	4	.94	
B x C x D	.01	4	.05	
A x B x C x D	.31	4	2.52	<.05
Error (Within Ss)	.12	256		

Appendix II

Summary of Analysis of Variance for MA Measured by the PPVT

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs Lo)	207.03	1	2.43	
B (Modality: Tact. vs Vis.)	5.62	1	.07	
A x B	.23	1	.00	
Error	85.05	36		

Appendix JJ

Summary of Analysis of Variance for CA

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs Lo)	5.63	1	.19	
B (Modality: Tact. vs Vis.)	7.22	1	.25	
A x B	9.03	1	.31	
Error	29.05	36		

Appendix KK

Summary of Analysis of Variance for PPVT Raw Score

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs Lo)	65.02	1	2.82	
B (Modality: Tact. vs Vis.)	4.22	1	.18	
A x B	.22	1	.01	
Error	23.04	36		

Appendix LL

Summary of
Analysis of Variance for IQ

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs Lo)	84.10	1	2.28	
B (Modality: Tact. vs Vis.)	.10	1	0	
A x B	.10	1	0	
Error	36.88	36		

Appendix MM

Summary of
Analysis of Variance for SES Index

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs Lo)	14478.03	1	194.43	< .001
B (Modality: Tact. vs Vis.)	24.02	1	.32	
A x B	34.23	1	.46	
Error	74.46	36		

Appendix NN

Summary of Analysis of Variance for
Errors in Stimulus Equivalence Judgements

Source	Mean Square	df	F Ratio	p Value
A (SES: Hi vs Lo)	1.25	1	.44	
B (Standard: Tact. vs Vis.)	162.45	1	57.79	< .005
A x B	6.05	1	2.15	
Error (Between)	2.81	36		
(Transfer: Intra vs Inter Modality)	92.45	1	39.81	< .005
A x C	4.05	1	1.74	
B x C	14.45	1	6.22	< .025
A x B x C	.45	1	.19	
Error (Within)	2.32	36		